Lessons from the field in the ELSA COVID-19 Substudy

Sample integrity and survey nonresponse in a mixed-mode design

NatCen
Social Research that works for society

Authors: Gianfranco Addario, Martin Wood
Date: 19/08/2022

At **NatCen Social Research** we believe that social research has the power to make life better. By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to put all our time and energy into delivering social research that works for society.

NatCen Social Research
35 Northampton Square
London EC1V 0AX
T 020 7250 1866
www.natcen.ac.uk

A Company Limited by Guarantee
Registered in England No.4392418.
A Charity registered in England and Wales (1091768) and Scotland (SC038454)
This project was carried out in compliance with ISO20252

Contents

Executive summary

The second wave of the ELSA COVID-19 Substudy

- The ELSA COVID-19 Substudy was conducted with members of ELSA, the English Longitudinal Study of Ageing. It collected data about the experiences of the population of England above the age of 52 in the first year of the COVID-19 outbreak.
- The study was carried out in two waves (summer and winter 2020). Its data can be analysed on its own, or with previous and future waves of the ELSA study.
- The second wave of the study was issued to 9,086 ELSA sample members and achieved a response rate of 74.8% (6,794 completed interviews).
- COVID-related restrictions meant ELSA's usual face-to-face interview approach could not be implemented. Instead, the Covid-19 waves used a sequential mixed-mode design; in the second wave, 83.2% of the interviews were completed on Web, while 16.8% on CATI (telephone interviewing).

Survey experiments

- Two experiments based around communication with sample members were carried out in the second wave of the ELSA COVID-19 Substudy.
- "The interview length experiment" assessed whether indicating in letters and emails a different interview length ("30 minutes" versus "up to 45 minutes") influenced the likelihood to take part in the study. We did not find statistically significant variations between the control and the experimental groups.
- "The charitable donation experiment" assessed whether communicating that it was possible to donate the face value of the conditional incentive to a charitable organisation increased the response rate. We did not find any statistically significant difference between those who were told about it and those who were not.
- We found that the decision of *making* a charitable donation was influenced by socio-economic elements and demographics, but also by having been told about this option (indicating the importance of communication material) and by having completed the interview on CATI (suggesting that it was partly driven by social desirability).

Nonresponse in the ELSA COVID-19 Substudy

- We analysed nonresponse between the standard ELSA study and the ELSA COVID-19 Substudy focusing on elements associated with survey nonresponse and nonresponse in a Web-CATI design. These were frequency of internet use, self-reported health conditions (general health, eyesight, hearing), level of education, age group, refreshment sample (cohort), sex, and ethnicity.
- We found that almost all subgroups considered in the analysis were less likely to take part in the substudy, compared to the previous CAPI waves; however, nonresponse was higher in some subgroups than in others. Categories associated with steeper drops in response rate were:
    - Accessing the internet less than once a month or never.
    - Having poor general health, poor eyesight, or poor hearing.
    - Being aged up to 65 or above 75.
    - Having joined the sample in recent refreshment cohorts (7 or 9).
- The response rate across subgroups in the substudy was lower compared to the CAPI waves, but not low in absolute terms. With the exceptions of sample

members with poor eyesight, all the subgroups considered in this analysis had a response rate above 60%.

- We also found that nonresponse was strongly mitigated by offering an offline interviewer-based mode (CATI) after the initial online self-administered mode (Web). The subgroups that performed better in a mixed-mode design, compared to the expected response rate from a unimodal Web design were:
    - Accessing the internet less than once a month or never.
    - Having poor eyesight.
    - Having achieved a level of formal education below university degree or equivalent.
    - Being above the age of 75.
    - Being part of the original ELSA sample (cohort 1) or of earlier refreshment samples (cohort 3, 4 or 6).
    - Being female.

Introduction

The ELSA COVID-19 Substudy was carried out using the sample of ELSA, the English Longitudinal Study of Ageing, during the first year of the COVID-19 outbreak in England. ELSA is a large-scale probability-based longitudinal study of the population of England above the age of 50, whose participants were recruited among those who took part in the Health Survey for England, along with their partners.

The ELSA COVID-19 Substudy was carried out in two waves of data collection. Fieldwork for the first wave was conducted between June and July 2020, while the second wave took place in November and December 2020.

This report offers an overview of the fieldwork for the second wave of the substudy, provides an assessment of methodological experiments that were carried out in the second wave, and draws conclusions about nonresponse across the entire ELSA COVID-19 Substudy. More detail of the general approach is provided in the report for the first wave of the substudy (Addario et al. 2020).

This report is divided in five different chapters. The first chapter presents fieldwork results from the second wave of the substudy and compares response rates and survey modes across both waves of the ELSA COVID-19 Substudy. The second chapter focuses on fieldwork design and, specifically, on how the approach diverged between the two waves of the substudy due to a slightly shorter fieldwork. The third and fourth chapters review two communication experiments that were carried out in the second wave, respectively the “interview length experiment” and the “charitable donation experiment”, describing their design, research questions and hypotheses, and results. The fifth chapter explores patterns of nonresponse between the ELSA main study and the ELSA COVID-19 Substudy, focusing on socio-economic elements and demographic characteristics known to be associated with nonresponse; it also assesses the role of the CATI mode in mitigating nonresponse bias associated with the Web mode.

1 Response rate and survey modes

This chapter offers an overview of the response rate and fieldwork outcomes of the second wave of the ELSA COVID-19 Substudy and of the modes used to complete the survey. The next few sections explore these dimensions and expand the discussion in a comparison between the first and the second wave of the study.

We found that the two waves of the ELSA COVID-19 Substudy had a comparable response rate and little difference, if any, was also observed for key subgroups of the sample. The split in survey modes was also similar between the two waves, with CATI being more common among those subgroups that seem to have been less engaged with the study.

1.1 Average response rate and fieldwork outcomes

The second wave of the ELSA COVID-19 Substudy achieved a final response rate of 74.8%, with 9,086 issued cases and 6,794 completed interviews. The mean response rate was similar to that achieved for the first wave of this study (75%)[1], despite the second wave having a shorter fieldwork period (6 weeks and 5 days) compared to the first wave (7 weeks and 5 days).

The unproductive cases at the second wave of the ELSA COVID-19 Substudy are almost equally split between unproductive due to no contact by the CATI interviewer (12.5%; this includes wrong or missing telephone numbers, rings out, call-blocking machines, fax and other forms of no contact) and unproductive after contact was made (12.8%, including refusals, either by proxy or in person, being unable to take part in the survey and other forms of refusal). It is important to remember that all the cases have been invited to take part online (either via email or via letter) ahead of telephone fieldwork, meaning that "no contact" only refers to attempts carried out by CATI interviewers (all cases are assumed to have received communication of the online survey). More details about the fieldwork outcomes for this second wave of the study are included in Table 1.

Table 1: ELSA COVID-19 W2 Substudy fieldwork outcomes

	n	%
Study sample	**9524**	
Found ineligible before fieldwork or study refusal	380	
Issued	**9144**	
Found ineligible during fieldwork	58	
Eligible for an interview at W2	**9086**	100.0%
Fully productive	**6794**	74.8%
No contact[1]	**1132**	12.5%
Unproductive after contact made	**1160**	12.8%
Refusal by proxy	68	
Refusal	817	

[1] 9,392 sample members were invited to take part in the first wave of the study, resulting in 7,040 completed surveys (Addario et al., 2020).

Ill or away during fieldwork period	123
Physically or mentally unable to take part	60
Other unproductive	92

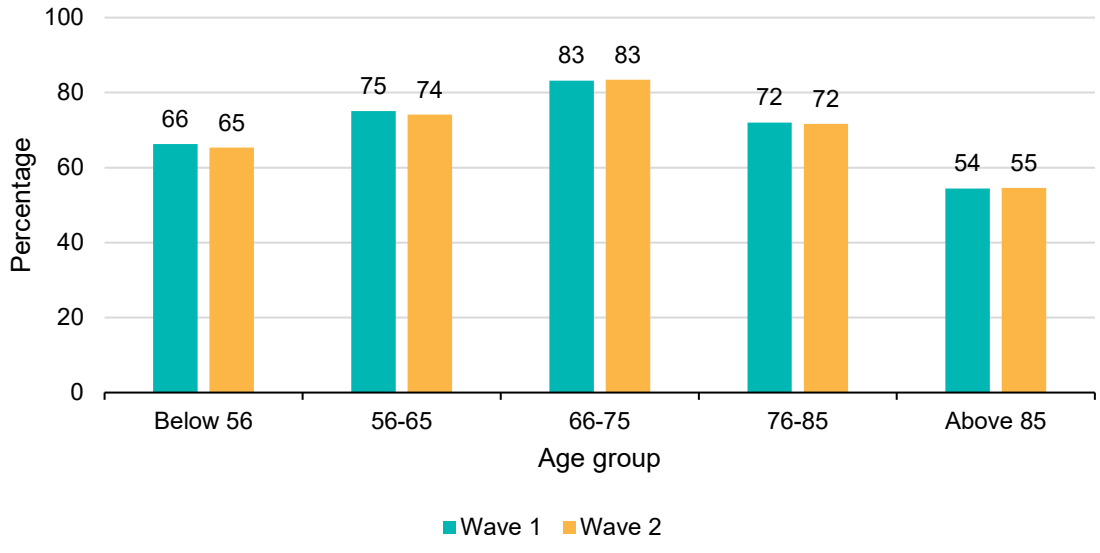
[1] Includes wrong number, rings out, call-blocking, fax and other forms of no contact

The longitudinal response rate, intended as the percentage of sample members eligible for an interview in the second wave of the study who took part in both the first and the second wave, was 92.4%. Only 15.5% of the 2,082 sample members who did not take part in the first wave of the study decided to complete the second wave.

1.2 Response rate by sample members' characteristics

The response rates by age groups show similar patterns in the two waves of the ELSA COVID-19 Substudy (Figure 1). In both waves it was at its highest, 83%, among sample members aged between 66 and 75, and at its lowest, about 54%, for sample members above the age of 85.

Figure 1: Response rate by age group in the first and second wave of the ELSA COVID-19 Substudy

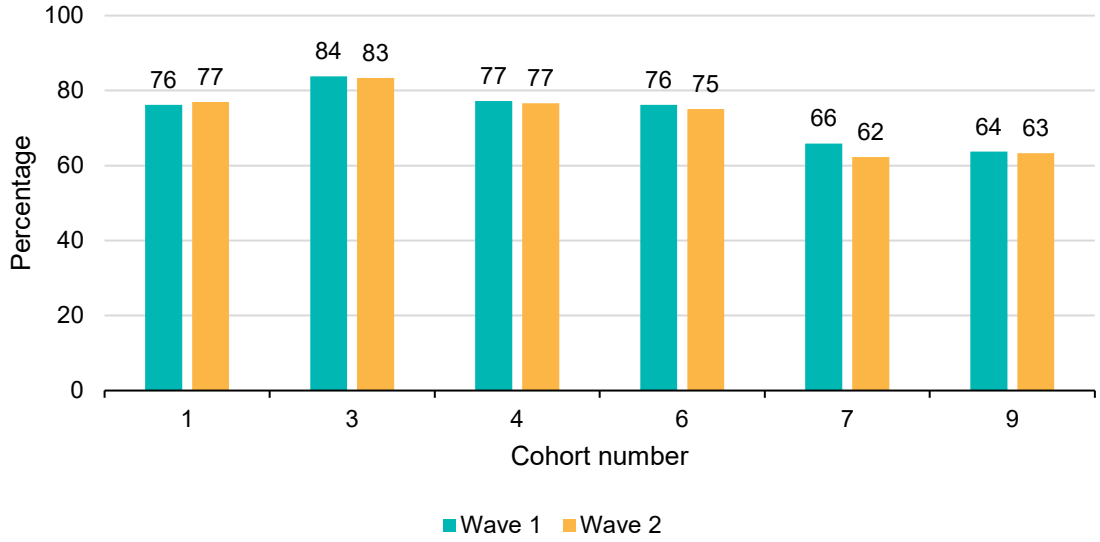


Wave 1 base: below 56 1,251, 56-65 2,332, 66-75 3,346, 76-85 1,899, above 85 564. Wave 2 base: below 56 1,248, 56-65 2,310, 66-75 3,292, 76-85 1,771, above 85 465.

Similar patterns between the two waves of the study were also found in the response rate by refreshment sample[2], or cohort (Figure 2). In both waves, it was lower amongst study participants who were invited to join the ELSA sample either in the 7th or the 9th wave of the main study and were therefore relatively new to it.

[2] The ELSA sample was refreshed 5 times in 9 waves. Additional sample members were recruited at wave 3, 4, 6, 7 and 9.

Figure 2: Response rate by cohort in the first and second wave of the ELSA COVID-19 Substudy



Wave 1 base: Cohort 1 4,106, Cohort 3 1,030, Cohort 4 1,614, Cohort 6 880, Cohort 7 393, Cohort 9 1,331. Wave 2 base: Cohort 1 3,879, Cohort 3 1,015, Cohort 4 1,562, Cohort 6 872, Cohort 7 390, Cohort 9 1,330.

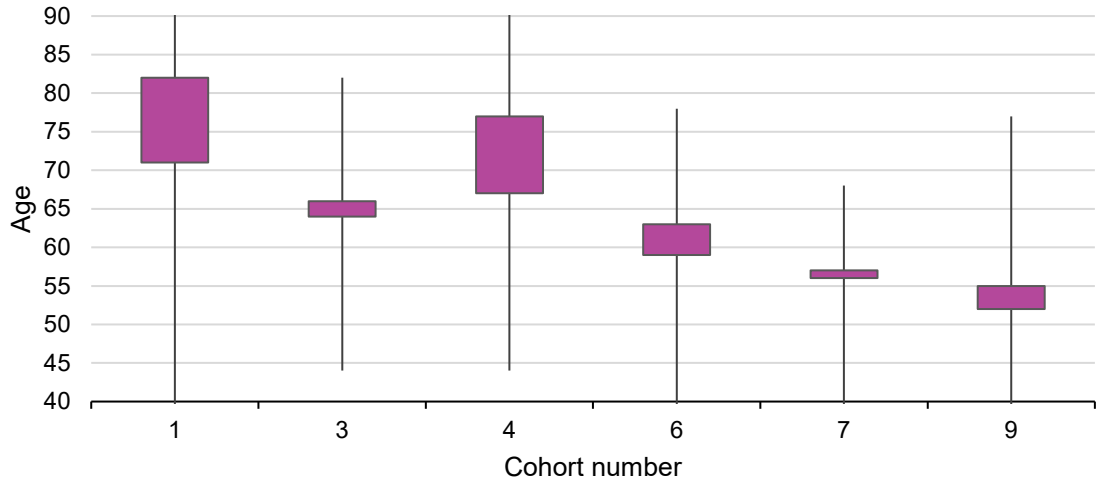
When interpreting these findings, it is important to note that the response rates by age and cohort are confounded, as shown by the age distribution in the different cohorts (Figure 3). Refreshment samples are regularly integrated in ELSA to compensate for the natural ageing of the existing sample, so that the study can still be representative of the English population aged 50 and above.

This means that cohort considerations in ELSA should be assessed factoring in the age element, as recent cohorts are generally younger than the earlier ones. Indeed, sample members in cohort 4 to 6 fell in the age groups associated with higher response rates, compared to cohort 1 (relatively older) and cohorts 7 and 9 (relatively younger).

While it is possible that the lower response rate amongst older sample members might be associated with the change of survey mode, cohort considerations might help explain the lower response rate between younger sample members. Lower response rates amongst the newest refreshment cohorts can also be found in the standard ELSA waves[3] and it is also largely confirmed by survey literature: earlier cohorts are found to be less vulnerable to longitudinal attrition (Jackle et al., 2015; Lugtig, 2014) than recent cohorts.

[3] The cooperation rate in ELSA W9 was 66% for Cohort 9, 80% for Cohort 6 and 7, 86% for Cohort 4, 89% for Cohort 3 and 88% for Cohort 1 (Pacchiotti et al., 2021).

Figure 3: Age distribution (interquartile range) by cohort in the ELSA COVID-19 sample (core members and their partners)

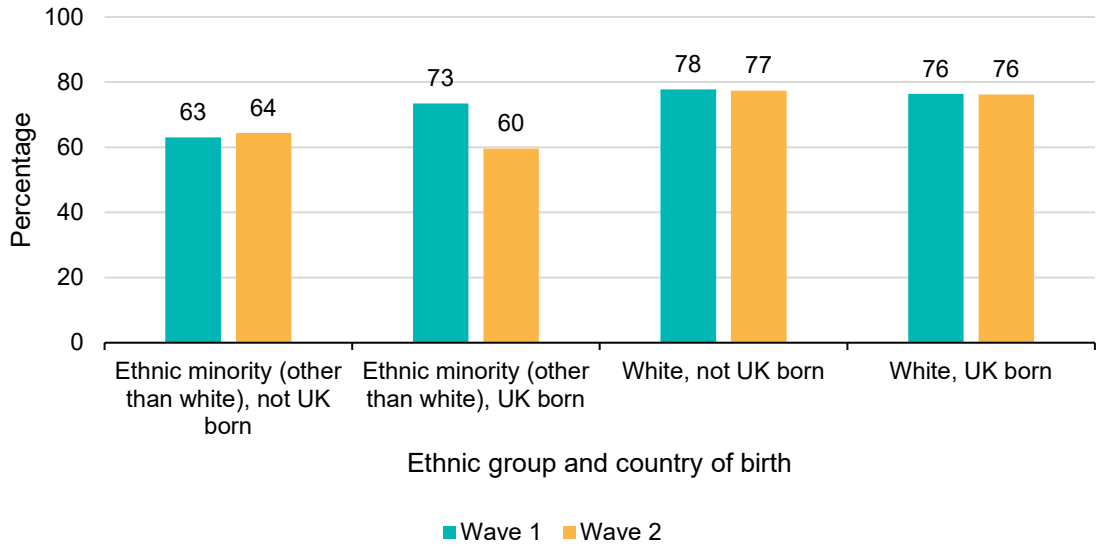


Wave 1 base: Cohort 1 4,106, Cohort 3 1,030, Cohort 4 1,614, Cohort 6 880, Cohort 7 393, Cohort 9 1,331. Wave 2 base: Cohort 1 3,879, Cohort 3 1,015, Cohort 4 1,562, Cohort 6 872, Cohort 7 390, Cohort 9 1,330.

Little differences in response rate between the two waves of the ELSA COVID-19 Substudy were also found when looking at ethnic identity and country of birth of the ELSA sample members (Figure 4).

Across the different groups, sample members with an ethnic minority background (not including white minorities) who were not born in the UK were less likely to take part in the study at the first wave, compared to the other groups.

Figure 4: Response rate by ethnic group, and by whether born in the UK, in the first and second wave of the ELSA COVID-19 Substudy

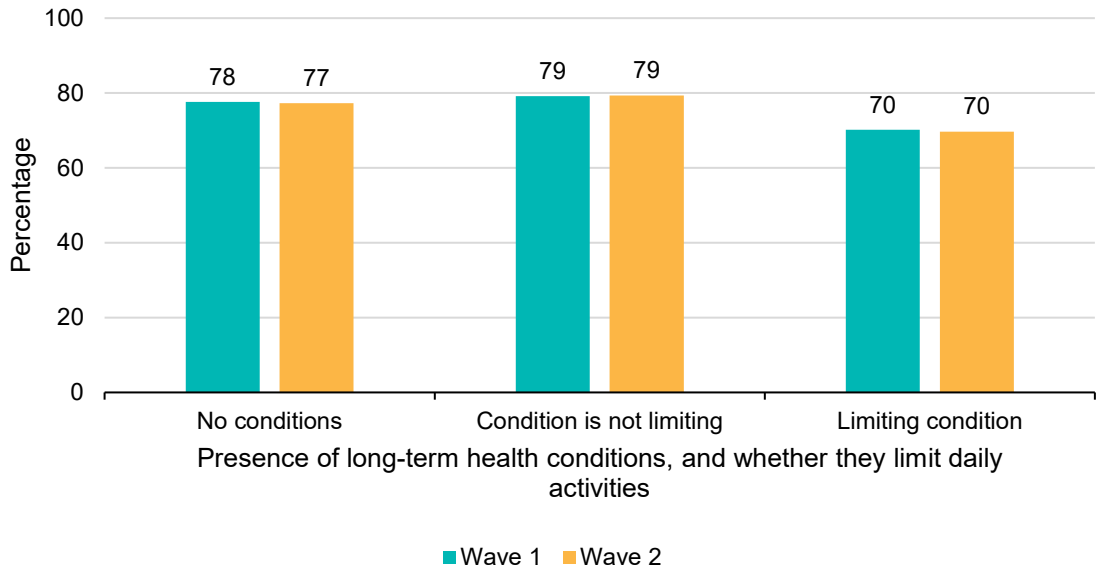


Wave 1 base: Ethnic minority (other than white) not UK born 376, Ethnic minority (other than white) UK born 94, White not UK born 707, White UK born 8,031. Wave 2 base: Ethnic minority (other than white) not UK born 368, Ethnic minority (other than white) UK born 94, White not UK born 687, White UK born 7,759.

Strong variations between the first and the second wave were found only across sample members with an ethnic minority background (not including white minorities) who were born in the UK. However, when interpreting these findings, it is important to remember that this group is particularly small (about 1% of the ELSA sample).

Survey participation rates were substantially the same between the first and the second wave of the study by presence of long-term health conditions and whether they limit daily activities (Figure 5). Sample members with limiting conditions were less likely to complete the study, compared to those with no conditions, or whose conditions did not limit their activities.

Figure 5: Response rate by presence of long-term conditions, and whether they limit daily activities, in the first and second wave of the ELSA COVID-19 Substudy

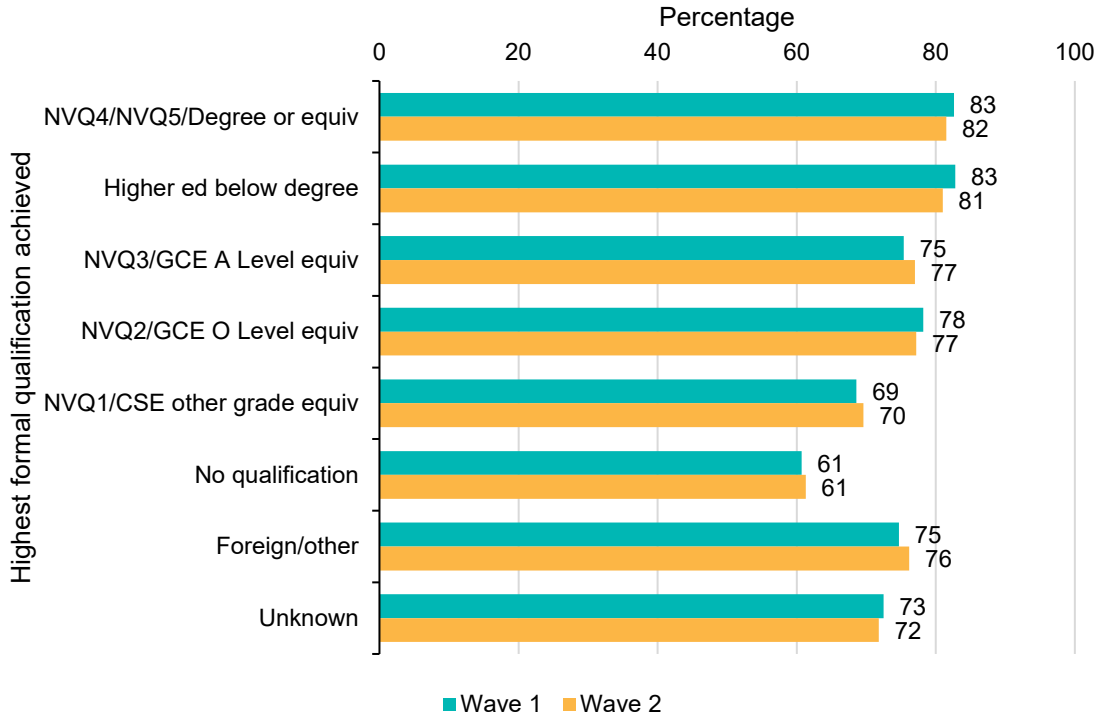


Wave 1 base: No conditions 4,236, Condition is not limiting 1,970, Limiting condition 3,068. Wave 2 base: No conditions 4,144, Condition is not limiting 1,927, Limiting condition 2,897.

Comparable patterns, with little variations, between the two waves was also observed for response rate by highest level of formal education achieved (Figure 6). In both waves, the likelihood of taking part was higher amongst sample members who achieved at least a university degree or equivalent, and decreased for lower levels of formal education, reaching the lowest levels amongst those who did not have any qualification.

This is also aligned to survey literature on nonresponse and longitudinal attrition, as higher levels of education are generally found to be associated with higher levels of cooperation (Watson and Wooden, 2009).

Figure 6: Response rate by highest formal qualification achieved in the first and second wave of the ELSA COVID-19 Substudy



Wave 1 base: NVQ4/NVQ5/Degree or equiv. 1,721, Higher ed below degree 1,167, NVQ3/GCE A Level equiv. 810, NVQ2/GCE O Level equiv. 1,723, NVQ1/CSE other grade equiv. 277, Foreign/other 767, No qualification 1,622, Unknown 1,305. Wave 2 base: NVQ4/NVQ5/Degree or equiv. 1,693, Higher ed below degree 1,129, NVQ3/GCE A Level equiv. 793, NVQ2/GCE O Level equiv. 1,683, NVQ1/CSE other grade equiv. 263, Foreign/other 741, No qualification 1,509, Unknown 1,275.

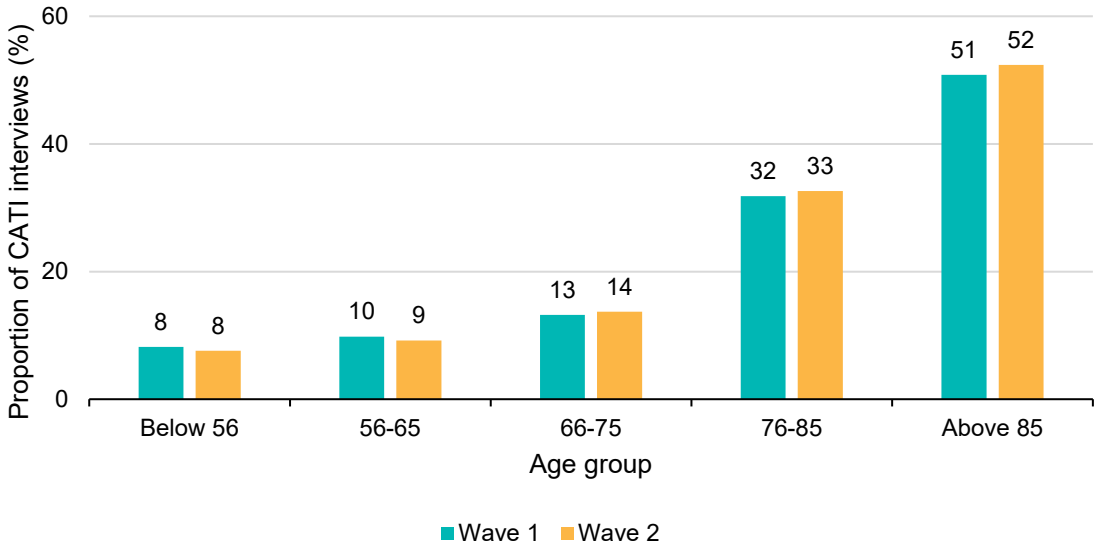
1.3 Interview modes

At the second wave of the ELSA COVID-19 Substudy, 83.2% of sample members who took part completed the questionnaire online, while 16.8% completed via CATI. This split is similar to that observed for the first wave, when 83.0% of the interviews were completed on Web and 17.0% on CATI.

The CATI mode was instrumental in the level of cooperation achieved amongst older sample members (Figure 7), as the proportion of survey interviews completed via CATI increased with age; while less than 1 in 10 survey interviews were completed via CATI among sample members below the age of 56, the proportion of CATI interviews was a little more than 1 in 2 among sample members above the age of 85.

The preference for the CATI mode amongst people in later life is likely to be associated to a range of socio-economic and health-related factors, such as low levels of IT literacy or health conditions that make it particularly difficult the use of IT devices. Data collected in the first wave of the ELSA COVID-19 Substudy indicates that 15% of the population of England aged 52 or above never use the internet, but the proportion increases to 57% between people above the age of 85. Only 17% of the people in this age group reported that they would like to use the internet more, and 78% of them listed the lack of sufficient IT skills as the reason for not using the internet more often.

Figure 7: Proportion of CATI interviews by age group in the first and second wave of the ELSA COVID-19 Substudy

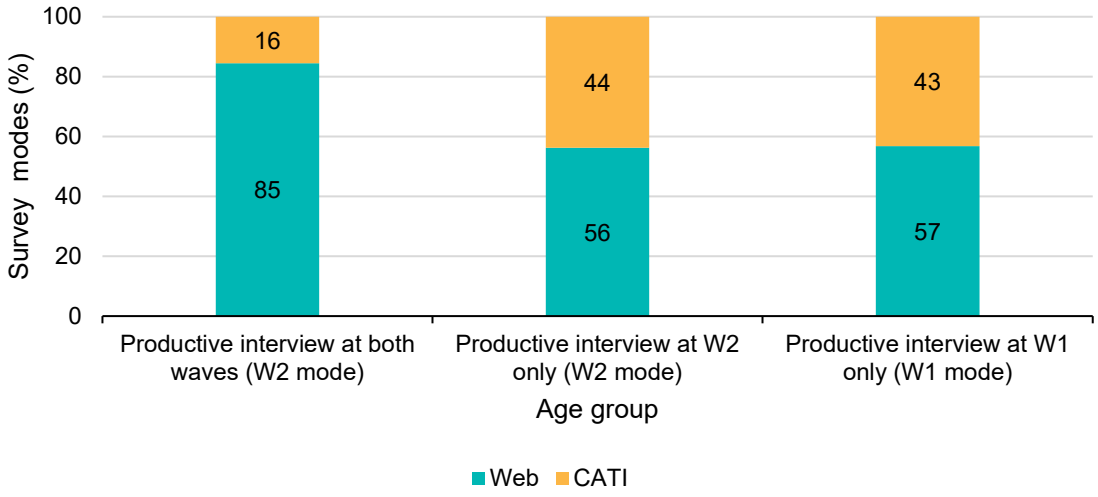


Wave 1 base: below 56 829, 56-65 1,752, 66-75 2,785, 76-85 1,367, above 85 307. Wave 2 base: below 56 815, 56-65 1,711, 66-75 2,744, 76-85 1,270, above 85 254.

The operationalisation of a CATI mode and telephone fieldwork was not only necessary to reach the 'offline' population, but also to actively increase the cooperation rate between sample members less engaged with the study (Calderwood et al., 2019).

As shown in Figure 8, 85% of the sample members who completed the first wave of the study took part in the second wave via Web, while the remaining 16% were interviewed on the telephone. The proportion of CATI interviews was greater (above 43%) amongst sample members who took part in only one wave (not in both waves), who were probably less engaged with the study. This suggests that a large number of interviews across less cooperative sample members would not have been achieved without telephone fieldwork.

Figure 8: Survey modes by participations in the first and second wave of the ELSA COVID-19 Substudy



Base: productive interview at both waves 6,472, productive interview at W2 only 322, productive interview at W1 only 558.

This proactive role of interviewers can also be understood looking at how survey modes changed between waves. Sample members who completed the second wave of on Web were particularly likely to have completed the first wave on Web too (96% had completed the first wave on Web, and 4% on CATI), suggesting that they did not need to be proactively convinced to take part in either wave.

The proportion of Web interviews at W1 amongest those who completed the second wave on CATI was larger (75% CATI and 25% Web). This suggestes that many sample members in this group might have had the skills or the possibility to take part on Web (as they did, in W1), but needed contact with an interviewer before they accepted to take part in the second wave.

2 A shorter fieldwork period

The fieldwork design of the ELSA COVID-19 Substudy was similar in both waves of the study. The key difference was a shorter fieldwork period in the second wave, which required some adustments to how the CATI fieldwork was structured and managed.

This chapter does not offer a comprehensive overview of the fieldwork design, which was already discussed in the methodological report at the end of the first wave (Addario et al. 2020). Instead, it focuses on the differences between the two waves, and assesses whether and how these differences have had an impact on the data collection results.

2.1 Fieldwork structure and compressed design

The fieldwork structure of the second wave of the ELSA COVID-19 Substudy was based on the same sequential mixed-mode approach with responsive CATI management operationalised in the first wave of the study. In practical terms, all sample members were firstly invited to complete a self-administered online questionnaire; those who didn't take part were later contacted by an interviewer and asked to complete the survey via telephone. CATI interviewers prioritised some cases over others, with those considered more unlikely to complete on Web receiving earlier contact and a greater investment of interviewer time and project resources.

The methodological aims of this fieldwork approach were the same in both waves (Addario et al. 2020: 14): reaching the offline population, budget optimisation through a web-first approach, and compensation of web nonresponse offering an interviewer-led mode.

There were, however, some variations to the design used in the first wave, primarily implemented to respond to the shorter fieldwork period for the second wave. The first wave of the study ran for 7 weeks and 5 days, while the second wave was originally planned to last 5 weeks and 5 days (from the 4th of November 2020 until the 13th of December 2020).

A shorter fieldwork period meant that the design needed to be compressed to achieve the same objectives. This was operationalised in three ways:

- An "early CATI group" consisting of sample members thought not to access the internet had been sent for CATI fieldwork from an early point in the first wave, but this was not implemented in the second wave fieldwork design.
- CATI fieldwork started earlier in the second wave (7 days after the beginning of Web fieldwork) compared to the first wave (12 days after the beginning Web fieldwork for the "Early CATI group" and 26 days after the beginning Web fieldwork for other sample members).
- Statistical modelling for the allocation of Web-unproductive sample members in CATI groups was carried out using W1 data and was not based on early fieldwork data (there was no sufficient time to collect data, analyse it and process the information for CATI fieldwork management between the start of Web fieldwork and the beginning of CATI fieldwork).

2.2 Modelling and CATI groups

The allocation of sample members in different groups for the management of CATI fieldwork followed the same principles of the first wave, with issued cases predicted to be more likely to complete on Web being released to telephone interviewers later during fieldwork, and cases with a stronger association with Web nonresponse receiving earlier interviewer contact. In both waves, the prioritisation was determined through the use of statistical modelling (binary logistic regression).

Nevertheless, the process of modelling the likelihood of Web completion diverged from the first wave. Statistical modelling could not be based on early fieldwork data, due to time limitations, so the model was built using both sample information and data collected in the first wave of the study:

- Sex
- Age
- Cohort
- Last interview
- Whether there was another ELSA sample member in the household
- Economic activity
- Index of Multiple Deprivation quintile
- Relationship status
- Frequency of internet use
- Whether use the internet to receive and send emails

The use of W1 data meant that these predictions could be made only for cases who had had a productive interview in the first wave of the study. Sample members who had been unproductive in the first wave were assigned the same probability of Web completion computed at the beginning of W1 fieldwork.

CATI groups for W2 fieldwork were consequently designed using this struture:

- Batch 1 included the 1,000 sample members (999 after post-fieldwork eligibility checks) who had had a productive interview in the first wave of the ELSA COVID-19 Substudy and had the lowest probability of Web completion (median 0.32). This is the only group that was allocated ahead of fieldwork and was meant to replace the “Early CATI” group of the first wave of the study.

Sample members who had not taken part in the study within two weeks from the beginning of Web fieldwork were allocated to three other batches. The allocation focused on minimising longitudinal nonresponse; sample members who had completed the first wave of the study were prioritised over those who did not take part in the first wave:

- Batch 2 included 1,252 cases, all with a productive interview in the first wave of the substudy. The median probability of Web completion was 0.89.
- Batch 3 was composed of 268 cases productive at W1 (median probability of Web completion 0.97) and 1,005 cases unproductive at W1 (median probability of Web completion 0.09).
- Batch 4 was formed of 269 cases productive at W1 (median probability of Web completion 0.99) and 1,012 cases unproductive at W1 (0.43).

Table 2: Summary of the CATI group design

CATI groups	Productive at W1		Unproductive at W1	
	n	Probability Web completion	n	Probability Web completion
Batch 1	999	0.32	0	-
Batch 2	1,252	0.89	0	-
Batch 3	268	0.97	1,005	0.09
Batch 4	269	0.99	1,012	0.43

2.3 Survey completion in CATI groups

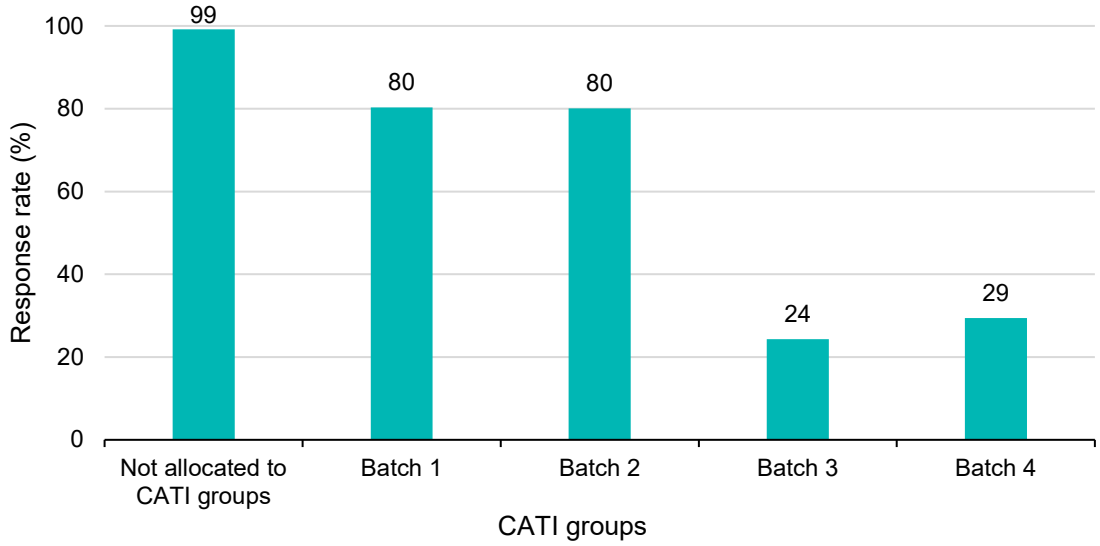
Batch 1, including 1,000 cases with a low likelihood of completing on Web, was released to telephone interviewers seven days after beginning Web fieldwork (11th of November). The second batch was released only after telephone contact was attempted, at least once, with all sample members from the first batch (19th of November).

On the 26th of November it was decided to extend the fieldwork of the second wave by one week (to 6 weeks and 5 days, ending on the 20th of December 2020 instead of the 13th) to ensure that CATI interviewers had sufficient time to interview all cases included in Batch 1 and 2; indeed, the participation rate across these batches was higher than anticipated and CATI interviewers needed more time before they could start working on Batch 3 and 4. These were released to telephone interviewers respectively on the 3rd of December and on the 10th of December.

The first two batches, those including sample members productive at the first wave and unlikely to complete on Web, achieved a final response rate of 80% (Figure 9), while it was lower in later batches (24% in Batch 3 and 29% in Batch 4). Batch 3 and 4 included a greater number of sample members who did not take part in the first wave of the Substudy and the CATI interviewers had less time available to encourage survey participation.

Response rate across those not allocated to CATI groups was 99%; however, it is important to remember that this group was mostly composed of sample members who completed on Web in the first two weeks of fieldwork, and an handful of cases who either did not want to be contacted by a CATI interviewer or had not provided a telephone number.

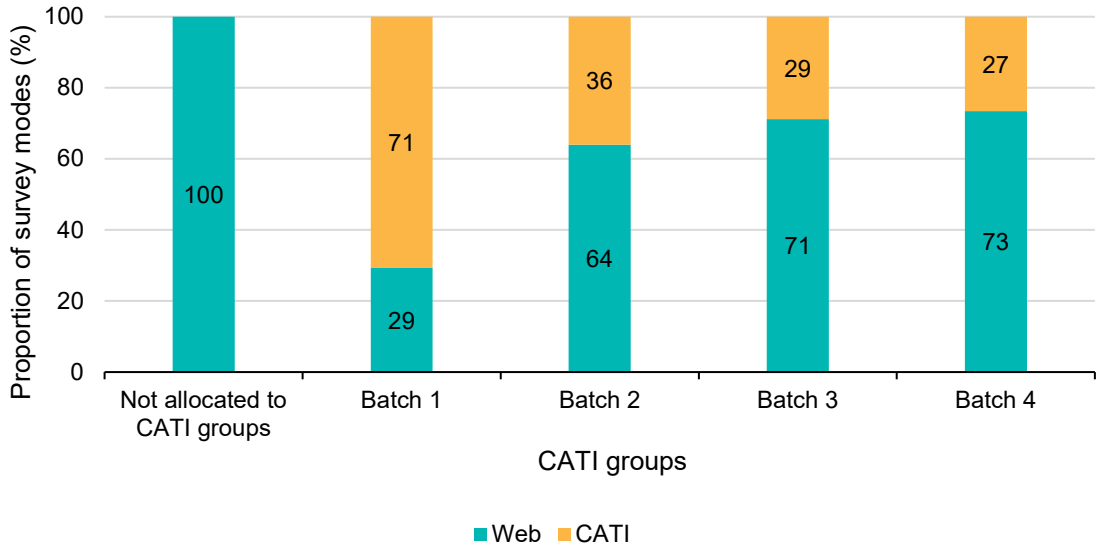
Figure 9: Response rate in the CATI groups



Base: Not allocated to CATI groups 4,339, Batch 1 999, Batch 2 1,252, Batch 3 1,273, Batch 4 1,281.

The difference in survey modes chosen to take part in the study between CATI groups reflects the modelling approach (earlier batches included cases less likely to complete on Web), but also the time available to CATI interviewers to work on the different batches (Figure 10). Batch 1 is the only one where more than half of the interviews were carried out on CATI, and the difference with Batch 2 is the result of the modelling approach, given that the CATI interviewers had sufficient time to work through both batches in full. More than 7 in 10 interviews in Batch 3 and 4 were completed on Web; this can be explained both by the modelling approach, but also by the fact that the CATI interviewers had less days to contact sample members in these groups.

Figure 10: Survey mode in the CATI groups



Base: Not allocated to CATI groups 4,304, Batch 1 802, Batch 2 1,003, Batch 3 309, Batch 4 376.

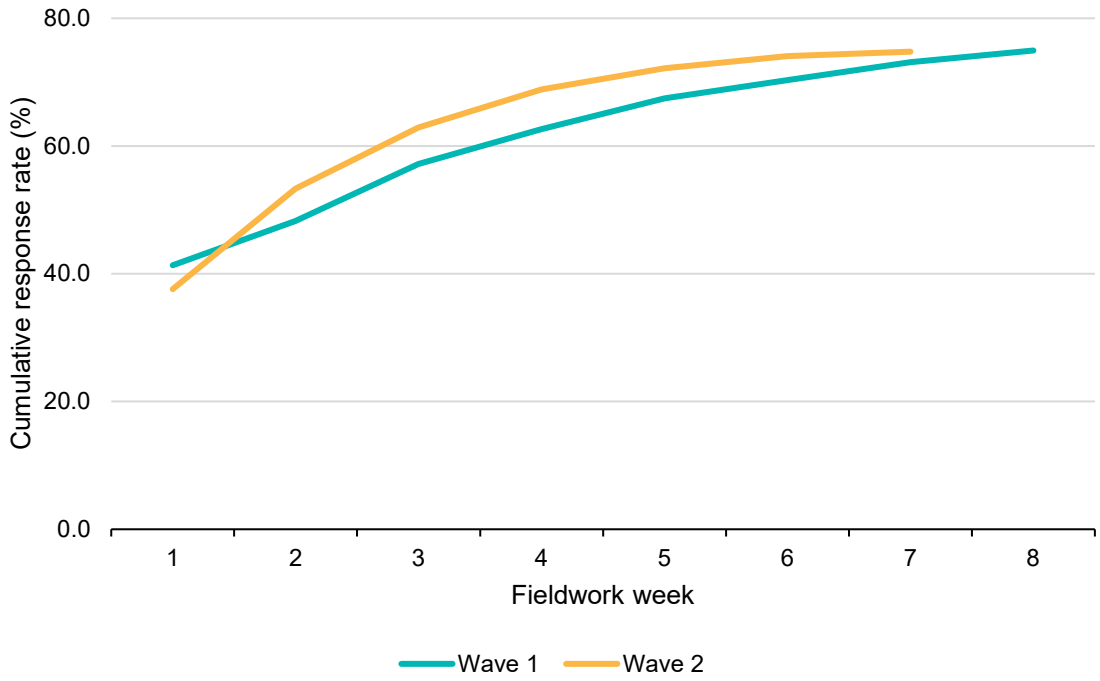
2.4 Pattern of survey completion

Despite differences in fieldwork design and the shorter fieldwork, findings from the previous chapter indicates that the mean response rate was consistent between the first and the second wave of the ELSA COVID-19 Substudy (75% in W1 and 74.8% in W2).

A comparison of patterns of survey completion between the two waves (Figure 11) shows that the response rate was similar in both the first and the last week of fieldwork, but the data collection in the second wave accelerated at a faster rate compared to the the first one. This is likely to be primarily a consequence of the earlier start of CATI fieldwork, but the faster rate might have also been influenced by the fact that this was the second wave of the study (sample members were already familiar with the design and knew what to do) and by seasonality (the wave took place in winter, not in the late sping/early summer months).

It is interesting that despite briging forward the beginning of CATI fieldwork, the balance of survey modes between the two waves is essentially unchanged, as discussed in the previous chapter. On the one hand this seems to confirm the validity of the CATI management approach followed in both waves: regardless of variations in the design, this approach ensured that the study budget was optimised and telephone interviewers did not establish contact too soon with sample members who could complete on Web. On the other thand, it might also be the result of a degree of stability in mode preference over time.

Figure 11: Cumulative response rate in the two waves



Base: Wave 1 9,524, Wave 2 9,144.

3 The interview length experiment

The second wave of the ELSA COVID-19 Substudy included a communication experiment on interview length. The aim of the experiment was to understand whether sample members informed of a longer interview length would have been less likely to engage with the study, compared to those who were informed of a shorter interview length. If researchers can be confident that longer interviews could be conducted without significant impact on response rates in the specific context of ELSA, this raise the possibility of significantly increasing the breadth and depth of detail to be collected in future. This chapter describes the structure and the findings of the experiment.

3.1 The design of the experiments

The experiment was design as a randomised control trial (RCT), where a random subset (25%) of the issued households (experimental group) was informed that the survey interview would have taken “up to 45 minutes”, while the remaining 75% (control group) were told that the interview would have taken “30 minutes”. This message was conveyed in initial letters and emails and in subsequent reminder letters and emails. The actual interview length was expected to be the same in both groups (30 minutes), given that both experimental group and control group were administered the same survey.

The allocation of sample members to experimental and control groups was carried out on a household level (not based on individuals) given that it was important to avoid conflicting messages about interview lengths to two or more sample members living at the same address. Sample members were assigned to the experimental group at random, within strata of the sample. The stratification was operationalised using derived variables associated with survey nonresponse:

- Whether the household included only one sample member, or more than one.
- Whether the average age of the sample members in the household was above or below 69.
- Whether at least one sample member in the household had a limiting health condition.
- Whether at least one sample member in the household achieved a qualification equivalent to, or above, university degree.
- Whether the household had been fully productive, partly productive, or unproductive in the first wave of the study.

The creation of the experimental group also took into account the presence of other experimental groups and sample splits, to ensure that all the allocations were balanced in the issued sample on a household level:

- Whether the household was assigned to another experiment (“The charitable donation experiment”, see Chapter 4).
- Whether the household was included in the first or in the second batch of invitation emails (invitation emails were sent in two batches for fieldwork operational reasons).

3.2 Research questions and hypotheses

This experiment was designed to answer a single research question that could inform sample members acceptability towards a slightly longer survey interview in a Web-CATI design:

(RQ1) Did a longer survey interview influence the decision to take part in the study?

Our hypothesis was that, given the ongoing engagement of sample members with the study and a particularly salient topic, a slightly longer interview length would not have made any difference in survey participation.

3.3 Results of the experiment

We found a small difference in the probability of taking part to the study between sample members allocated to the experimental group, compared to those in the control group. However, this difference was not statistically significant.

Table 3: Logistic regression models – interview length experiment and likelihood to take part in the study

	Model 1		**Model 2**		**Model 3**	
	Coef	*Se*	*Coef*	*Se*	*Coef*	*Se*
Intercept	1.10*	0.02	1.88*	0.14	-0.95*	0.19
Experimental group	-0.06	0.05	-0.05	0.06	-0.14	0.08
Cohort (ref: 1-6)						
Cohort 7-9			-0.64*	0.11	-0.54*	0.15
Age (ref: 55 or below)						
Age 56-65			-0.02	0.11	-0.08	0.15
Age 66-75			0.43*	0.13	0.19	0.18
Age 76-85			0.01	0.14	-0.35	0.19
Age above 85			-0.54*	0.16	-0.88*	0.22
Education level (ref: degree or above)						
No qualification			-0.68*	0.09	-0.30*	0.13
Qualification below degree			-0.09	0.07	0.01	0.11
Unkown			-0.22*	-0.10	-0.23	0.14
Frequency internet use (ref: almost every day or more)						
At least once a week or less			-1.03*	0.06	-0.72*	0.09
Unknown			-3.03*	0.13	-0.70*	0.16
Productive wave 1 interview (ref: yes)						
No					3.98*	0.08
Sample size	*9,086*		*9,048*		*9,048*	
AIC	*10,266*		*8,713*		*5,338*	
Pseudo-R² (McFadden)	*0.0*		*0.15*		*0.48*	

Dependent variable: having taken part in the second wave of the ELSA COVID-19 Substudy (productive survey interview = 1; unproductive interview = 0)
Note: Coefficients reported as log-odds.
* Statistically significant (alfa threshold 0.05).

The difference between these two groups was still not significant after controlling for other variables associated with survey nonresponse, such as cohort, age, education level and frequency of using the internet. We did not find any differences even after expanding the analsys to include the outcome at the previous waves (whether the sample member had a productive interview at W1 of the substudy).

The effect of the experiment did not change for subgroups, tested fitting interaction terms to the model. We did not find any statistically significant variation of the impact of the experiment by having taken part in the first wave of the study, age groups, cohort, frequency of internet use and level of education.

3.4 Conclusions of the experiment

The experiment on interview length confirmed the initial hypothesis. Small variations in the interview length do not seem to be an important driver in the decision to take part in a follow-up study for sample members of longitudinal panel studies.

Although we suspect this is likely to be driven by the ongoing commitment and relationship that sample members have developed with the study, it is also true that we did not find any statistical evidence to infer that the effect was different for more recent cohorts (those who joined the study between Wave 7 and Wave 9). These findings provide reassurance that, at least for salient topics, questionnaire content can be increased relatively significantly in web-CATI approaches with little impact on response.

4 The charitable donation experiment

The second wave of the ELSA COVID-19 Substudy included a second communication experiment, focusing on the role of incentives on response rate. As with the first wave, a conditional incentive was offered to all sample members with a productive interview, but the second wave also gave the possibility to offer the face value of the incentive to a charitable organisation.

Not all the sample members were informed about this possibility upfront, so that the effect on response rate of implementing this new strategy could be evaluated.

4.1 The design of the experiments

A random subset of 50% of the households issued to the study were informed that it would have been possible to donate the face value of the incentive (£10) to a charitable organisation (experimental group) or to receive a shopping voucher, while the remaining 50% of the households were only informed that they would have received a shopping voucher (£10) at the end of the interview (control group).

Subsequently, the option of making a charitable donation was offered to all sample members who completed the survey. Those who preferred this option could choose between four different charitable organizations (Age UK, British Heart Foundation, Action for Children, and The National Emergency Trust Coronavirus Appeal).

At the end of the study, a 'thank you' letter was sent to all the sample members who opted for a charitable donation, listing the total amount that was raised and donated to the charity they had chosen.

The allocation of sample members to experimental and control groups used the same system of randomisation and sample stratification described for the "Interview length experiment" (see Chapter 3).

4.2 Research questions and hypotheses

There were four research questions that we wanted to answer with this experiment:

(RQ1) Was the likelihood to take part in the study influenced by being given the option of making a charitable donation?

The option of making a charitable donation at the end of the study might have been appealing for some subgroups of respondents and, consequently, increased their cooperation in the study. However, given the high response rate in the second wave of the ELSA COVID-19 Substudy, we expected that having been informed about the possibility to make a charitable donation at the end of the interview did not have any effect on the likelihood to take part in the study (H1).

(RQ2) Was the likelihood to donate the incentive to charity influenced by having been pre-informed about this opportunity?

We expected that having been informed about the possibility to make a charitable donation did not have any impact on the decisions of whether to make, or not, a

charitable donation at the end of the interview (H2). In other terms, we expected a similar proportion of charitable donations in the control and in the experimental group.

(RQ3) Was the likelihood to donate the incentive to charity driven by key demographics and personal characteristics?

Our hypothesis (H3) was that we would have seen a greater preference for a charitable donation over shopping vouchers among sample members with a greater level of education and amongst the older age groups. We expected this to be driven by a mix of altruistic reasons (such as social responsibility and willingness to "give back") and personal reasons (including cost-opportunity considerations on how difficult, or easy, would have been to redeem and use the voucher).

(RQ4) Was the likelihood to donate the incentive to charity associated with the voucher preference expressed in in the first wave of the study?

At the end of the first wave of the study, sample members were not offer the possibility to opt for a charitable donation, but they were given the opportunity to opt-out from receiving a voucher. The "No voucher" option was selected by 5% of the sample members who completed a survey interview in the first wave of the substudy, while 95% opted for either a postal or a digital voucher. In this second wave we expected that sample members who opted for "No voucher" in the first wave of the study were more likely to have opted for a charitable donation at the end of the second wave (H4).

4.3 Effect on the likelihood to take part in the study

Previous communication about the possibility to make a charity donation at the end of the survey interview did not influence the likelihood to take part in the study. Indeed, in line with the first hypothesis (H1), the data suggested that sample members in the control group and those in the experimental group had an equal probability to take part in the study (the difference between the two groups was very small and not statistically significant).

This was true also after controlling for dimensions expected to be associated with survey participation (such as cohort, age, level of education, frequency of internet use and having participated in the first wave of the study).

Table 4: Logistic regression models – charitable donation experiment and likelihood to take part in the study

	Model 1		**Model 2**		**Model 3**	
	Coef	*Se*	*Coef*	*Se*	*Coef*	*Se*
Intercept	1.10*	0.034	1.88*	0.14	-0.97*	0.19
Experimental group	-0.03	0.05	-0.02	0.05	-0.04	0.07
Cohort (ref: 1-6)						
Cohort 7-9			-0.64*	0.11	-0.53*	0.15
Age (ref: 55 or below)						
Age 56-65			-0.02	0.11	-0.09	0.15

Age 66-75			0.43*	0.13	0.20	0.18
Age 76-85			0.01	0.14	-0.34	0.19
Age above 85			-0.54*	0.16	-0.88*	0.22
Education level (ref: degree or above)						
No qualification			-0.68*	0.09	-0.30*	0.13
Qualification below degree			-0.09	0.07	0.01	0.11
Unkown			-0.22*	-0.10	-0.23	0.14
Frequency internet use (ref: almost every day or more)						
At least once a week or less			-1.03*	0.06	-0.72*	0.09
Unknown			-3.03*	0.13	-0.71*	0.16
Productive wave 1 interview (ref: yes)						
No					3.98*	0.08
Sample size	*9,086*		*9,048*		*9,048*	
AIC	*10,267*		*8,714*		*5,341*	
Pseudo-R² (McFadden)	*0.0*		*0.15*		*0.48*	

Dependent variable: having taken part in the second wave of the ELSA COVID-19 Substudy (productive survey interview = 1; unproductive interview = 0)
Note: Coefficients reported as log-odds.
* Statistically significant (alfa threshold 0.05).

The effect of having received a communication about the possibility to make a charitable donation did not really change for subgroups of the sample (productive interview at the previous wave, age, internet and level of education). These were tested fitting interaction terms to the model.

4.4 Likelihood of making a charity donation

Contrarily to our second hypothesis (H2), the likelihood of opting for a charity donation at the end of the interview, in place of a digital or physical shopping voucher, was influenced by having been informed about this possibility in the study letters and emails. After controlling for age, cohort effects, level of education and frequency of internet use, we found sufficient statistical evidence to infer that sample members who had been assigned to the experimental group were more likely to have opted for a charity donation, than those assigned to the control group (and, therefore, not informed until the end of their survey interview about the possibility to donate the face value of the shopping voucher to a charitable organisation).

Table 5: Logistic regression models – likelihood to make a charitable donation

	Model 1		**Model 2**		**Model 3**	
	Coef	*Se*	*Coef*	*Se*	*Coef*	*Se*
Intercept	-1.26*	0.04	-0.73*	0.18	-0.05	0.20
Experimental group	0.15*	0.06	0.16*	0.06	0.15*	0.06
Cohort (ref: 1-6)						
Cohort 7-9			-0.63*	0.15	-0.60*	0.15
Age (ref: 55 or below)						
Age 56-65			-0.20	0.16	-0.16	0.16
Age 66-75			0.15	0.17	0.19	0.17
Age 76-85			0.53*	0.18	0.50*	0.18

Age above 85			0.67*	0.22	0.56*	0.22
Education level (ref: degree or above)						
No qualification			-1.15*	0.11	-1.24*	0.12
Qualification below degree			-0.60*	0.07	-0.63*	0.07
Unkown			-0.81*	0.11	-0.81*	0.11
Frequency internet use (ref: almost every day or more)						
At least once a week or less			-0.22*	0.08	-0.53*	0.09
Unknown			0.18	0.30	0.13	0.30
Interview mode (ref: CATI)						
Web interview					-0.74*	0.09
Sample size	*6,791*		*6,766*		*6,766*	
AIC	*7,393*		*7,110*		*7,043*	
Pseudo-R² (McFadden)	*0.0*		*0.04*		*0.05*	

Dependent variable: having made a charitable donation (yes = 1; no = 0)
Note: Coefficients reported as log-odds.
* Statistically significant (alfa threshold 0.05).

Sample members' socio-economic demographics also were associated with greater likelihood of opting for a charitable donation at the end of the survey. Variations were seen across age groups and level of educational attainment. While these findings related to age and level of education were aligned to our third hypothesis (H3), we also found the likelihood of making a charitable donation associated with cohort and frequency of internet use.

Interestingly, the mode of survey completion also heavily influenced the likelihood of opting for a charity donation. The probability of opting for a shopping voucher was higher among those who took part in Web, compared to those who completed a CATI interview.

4.5 Incentive preferences between the first and the second wave of the substudy

The option to make a charity donation was particularly common among those who had refused to receive a voucher at the end of the first wave interview, in line with our hypothesis (H4). 84% of those who did not want a voucher at W1 decided to make a charity donation, compared to 26% of those who had requested a digital shopping voucher and 18% of those who opted for a physical shopping voucher.

4.6 Conclusions of the experiment

The substantial independence between having been informed about the possibility to make a charitable donation and the decision to take part in the study (RQ1) can be explained with the same reasons discussed in for the "Interview length experiment" (Chapter 3). ELSA sample members are likely to be highly engaged with the study, and the possibility to make a donation to a charitable organisation at the end of the study might be well received (and therefore have some bearing on longer-term commitment to the study) but does not necessarily influence their levels of cooperation in the current wave.

Similarly, the association between sample members' demographics and the likelihood of opting for a charitable donation at the end of the second wave was somehow expected. The most interesting elements we found in the data were the variations of the likelihood of making a charitable donation between the "control group" and the "experimental group" and the variations by survey mode.

The differences between the two experimental groups highlight the importance of the information presented to sample members in the survey communication material. A few words about the incentives were able to direct a statistically significant number of sample members to prefer a charitable donation over a shopping voucher; this suggests that the communication material might be equally able to influence substantial answers to survey questions and the content of the communications should be drafted carefully to avoid any induced bias.

Variations in the likelihood of opting for a charitable donation across survey modes can probably be explained with mode effects and, primarily, by elements of social desirability. Sample members giving their preference to a telephone interviewer were more likely to opt for the most "socially responsible" or "socially acceptable" option (that is: making a charitable donation), while sample members taking part on a self-administered mode (without interviewer effects) were more likely to make a choice considering their own advantage (such as receiving a shopping voucher) without fearing the silent disapproval of an interviewer.

Finally, the relationship between the voucher options in the first wave and in the second wave suggested that the opportunity to make a charitable donation was particularly appealing among those who had opted out from a shopping voucher in the first wave and who were probably not so interested in receiving one anyway. Among those who refused a voucher in the first wave, 84% opted for a donation, 8% confirmed that they did not want a voucher and did not want to make a donation and 9% asked for a shopping voucher. In absolute numbers, however, most of the charitable donations were made by sample members who had requested a shopping voucher in the first wave, given that they were the vast majority (considering only those who completed both waves, 6,150 sample members had requested a shopping voucher at the end of the first wave, while only 321 had refused to receive one).

5 Changes in nonresponse between the CAPI and the Web-CATI designs

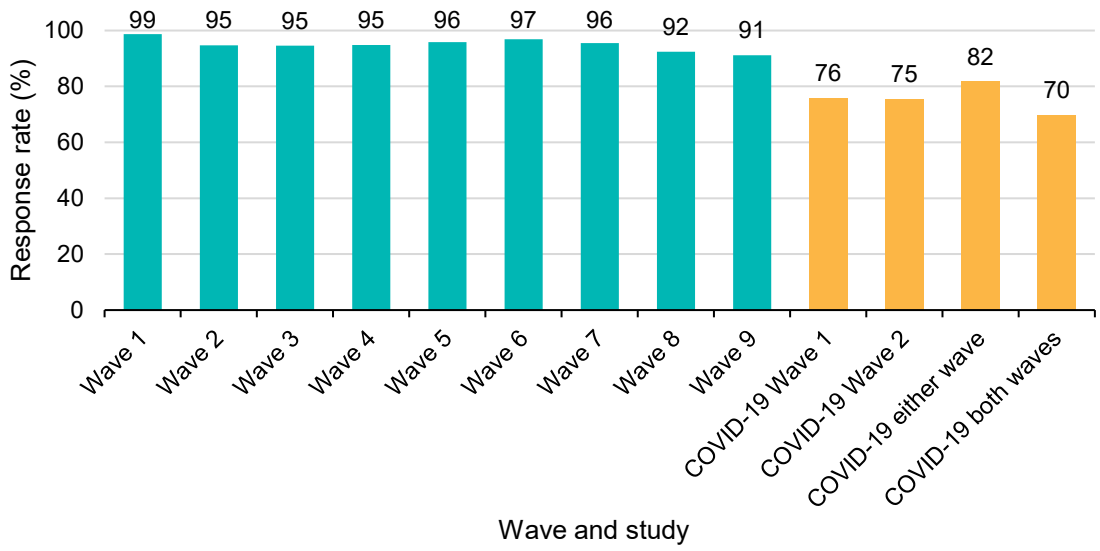
This chapter offers an assessment of nonresponse in the ELSA COVID-19 Substudy for core members of the sample. It is divided in two parts: the first part explores how survey participation has changed across the regular ELSA study and the two waves of the substudy, with a focus on socio-economic elements and demographic characteristics which are known to be associated with survey nonresponse; the second part of the chapter will explore the role played by the sequential mixed-mode design in increasing response rate and mitigating nonresponse bias from the Web survey mode.

Differently from the analysis carried out in other chapters of this report, which encompass all cases issued to the substudy, the analysis in this chapter is based on core members only and excludes their partners.

5.1 Response rate the ELSA

The response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy has always been above 90%, with a small decrease in the most recent waves (8 and 9), following the integration of refreshment samples at wave 7 and wave 9 (Figure 12).

Figure 12: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy



Base: 7,689.

The response rate in the ELSA COVID-19 Substudy for core members was lower, compared to the regular CAPI waves, but not low in absolute terms. Both waves concluded with a response rate above 75% among core sample members of the ELSA study eligible in the first wave of the ELSA COVID-19 Substudy.

5.2 Nonresponse in the ELSA COVID-19 Substudy

The first research question of this chapter aims at understanding how the achieved sample composition has changed between the ELSA study and the ELSA COVID-19 Substudy. In other words, we want to understand who did not take part to the study under this new design, which involved a change in the frequency and length of the interview and a shift of survey mode.

(RQ1) How did participation in the ELSA study changed between the CAPI and the Web-CATI design?

We take an exploratory approach, looking at changes across socio-economic elements and demographics that are either known to be associated with nonresponse, or are expected to be negatively correlated with survey participation in a Web-CATI design. Examples of the former group of variables are sex, ethnicity, educational attainment, and general health, while the latter group includes frequency of internet use, level of eyesight and level of hearing.

Statistical testing was conducted through multilevel models for change; each model tested whether the association between the variable of interest and the likelihood of taking part in the study statistically varied over time, and did not include additional control variables. This means that findings presented should be interpreted considering the potential confounded effect of age and cohort.

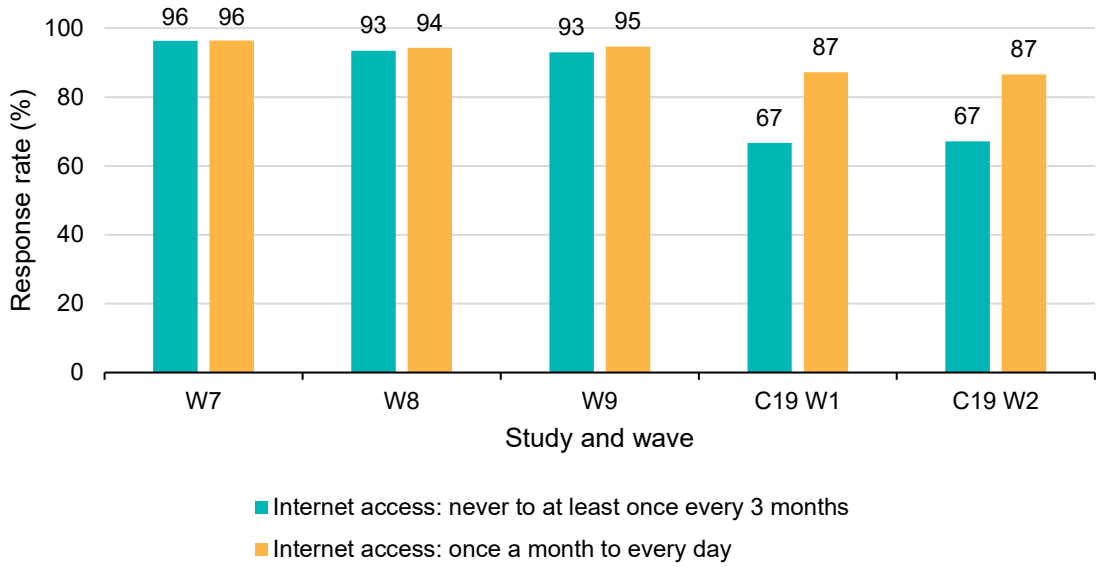
5.2.1 Frequency of internet use

Computer skills and IT literacy is a prime element of nonresponse in Web fieldwork, which is often paired in mixed-mode designs with offline modes, such as CATI, to ensure that the offline population is covered in the study (Becker, 2021).

In our analysis we used 1-wave lagged variables (internet use recorded at the previous wave, t-1, to assess survey participation in the following wave, t)[4]. We found evidence that, despite the use of a mixed-mode approach, sample members who accessed the internet less frequently (at least once in three months) or never, were less likely to have taken part in the ELSA COVID-19 Substudy compared to those who accessed the internet monthly or more frequently (Figure 13). Differences between these two groups of internet users were not statistically significant in previous CAPI waves.

[4] Questions about internet use were asked in ELSA after wave 6, therefore lagged variables can be used in the analysis from wave 7 onwards. The information for both waves of the COVID-19 Substudy were taken from the ninth CAPI wave.

Figure 13: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by frequency of internet use



Base: one a month to everyday 4,154 (W7), 4,366 (W8), 4,331 (W9), 5,007 (C19 W1), 4,927 (C19 W2); never to at least once every three months 1,508 (W7), 1,330 (W8), 1,199 (W9), 1,115 (C19 W1), 1,1016 (C19 W2).

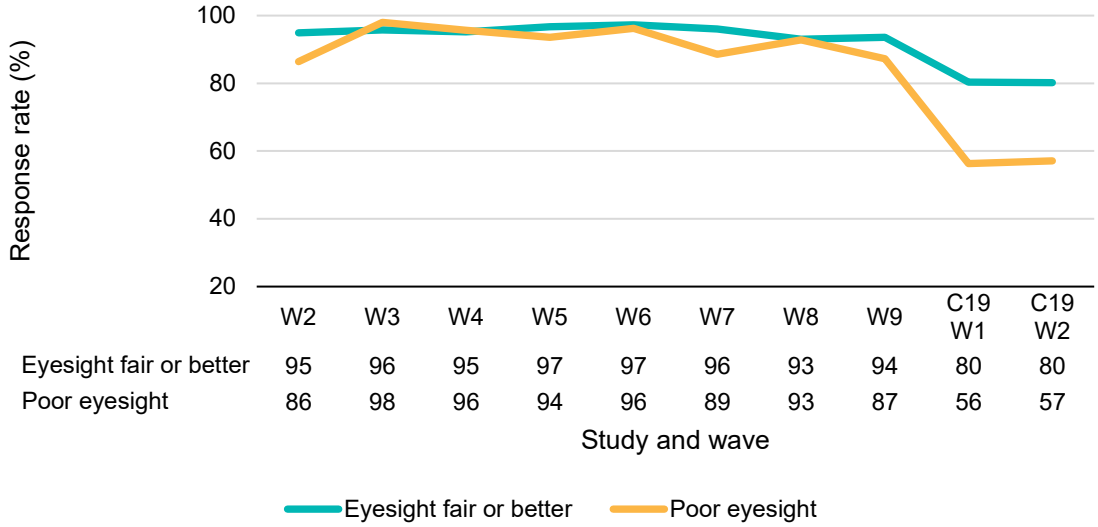
5.2.2 Self-reported health conditions

Health conditions are found to be a predictor of nonresponse in surveys, especially when they do not use in-person interviewing (Young et al., 2006). We focused on three types of self-reported health conditions:

- Level of eyesight (important for completion on Web, as participants need to read the self-administered questionnaire on the screen of an IT device).
- Level of hearing (important for completion on CATI, as participants need to maintain a phone conversation with an interviewer without the aid of visual cues).
- Level of general health.

Using 1-wave lagged health variables (data from both waves of the COVID-19 Substudy were taken from ELSA W9), we found that participants who reported poor eyesight, or poor hearing, or poor general health, were predicted a steeper drop in the probability of taking part in the COVID-19 Substudy, compared to sample members who reported better health outcomes. Poor health outcomes did not seem to predict higher nonresponse in the CAPI waves.

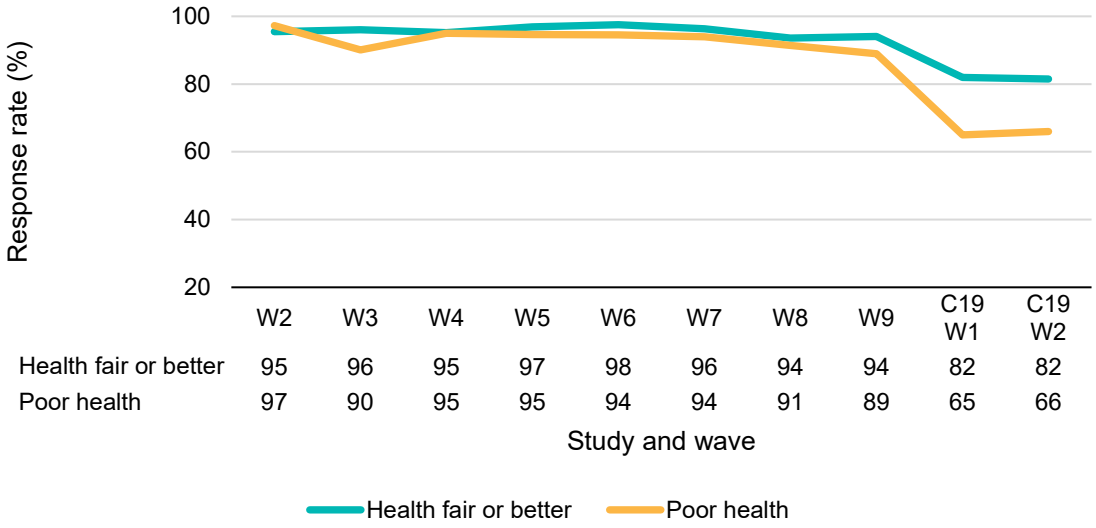
Figure 14: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by level of eyesight



	W2	W3	W4	W5	W6	W7	W8	W9	C19 W1	C19 W2
Eyesight fair or better	95	96	95	97	97	96	93	94	80	80
Poor eyesight	86	98	96	94	96	89	93	87	56	57

Base: Eyesight fair or better: 3,726 (W2), 3,541 (W3), 4,135 (W4), 5,496 (W5), 5,575 (W6), 6,204 (W7), 6,377 (W8), 6,108 (W9), 6,839 (C19 W1), 6,614 (C19 W2); Poor eyesight: 59 (W2), 50 (W3), 69 (W4), 95 (W5), 106 (W6), 131 (W7), 125 (W8), 181 (W9), 206 (C19 W1), 189 (C19 W2).

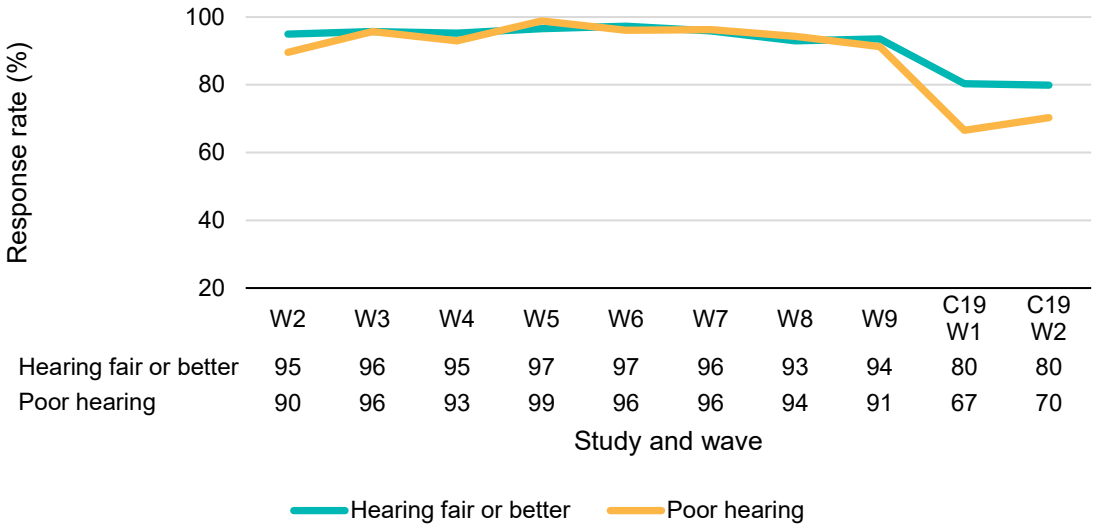
Figure 15: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by level of general health



	W2	W3	W4	W5	W6	W7	W8	W9	C19 W1	C19 W2
Health fair or better	95	96	95	97	98	96	94	94	82	82
Poor health	97	90	95	95	94	94	91	89	65	66

Base: Health fair or better: 1,796 (W2), 3,450 (W3), 4,028 (W4), 5,275 (W5), 5,307 (W6), 5,853 (W7), 5,991 (W8), 5,722 (W9), 6,326 (C19 W1), 6,141 (C19 W2); Poor health: 73 (W2), 133 (W3), 161 (W4), 245 (W5), 272 (W6), 348 (W7), 348 (W8), 416 (W9), 520 (C19 W1), 479 (C19 W2).

Figure 16: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by level of hearing

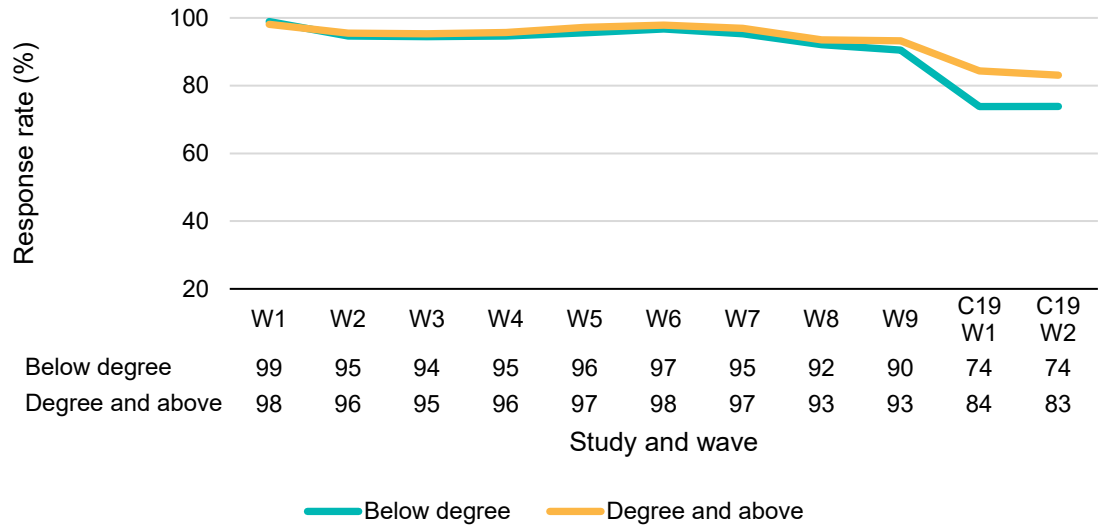


Base: Hearing fair or better: 3,669 (W2), 3,473 (W3), 4,075 (W4), 5,415 (W5), 5,478 (W6), 6,090 (W7), 6,241 (W8), 5,991 (W9), 6,717 (C19 W1), 6,511 (C19 W2); Poor hearing: 115 (W2), 117 (W3), 129 (W4), 176 (W5), 203 (W6), 244 (W7), 264 (W8), 299 (W9), 329 (C19 W1), 293 (C19 W2).

5.2.3 Highest level of formal education

We found differences in the likelihood of taking part in the ELSA COVID-19 Study by higher level of formal education. Those who had a level of education below university degree or equivalent were less likely to take part in the ELSA COVID-19 Substudy, compared to those with a level of education equal or above university degree, or equivalent.

Figure 17: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by level of formal education achieved



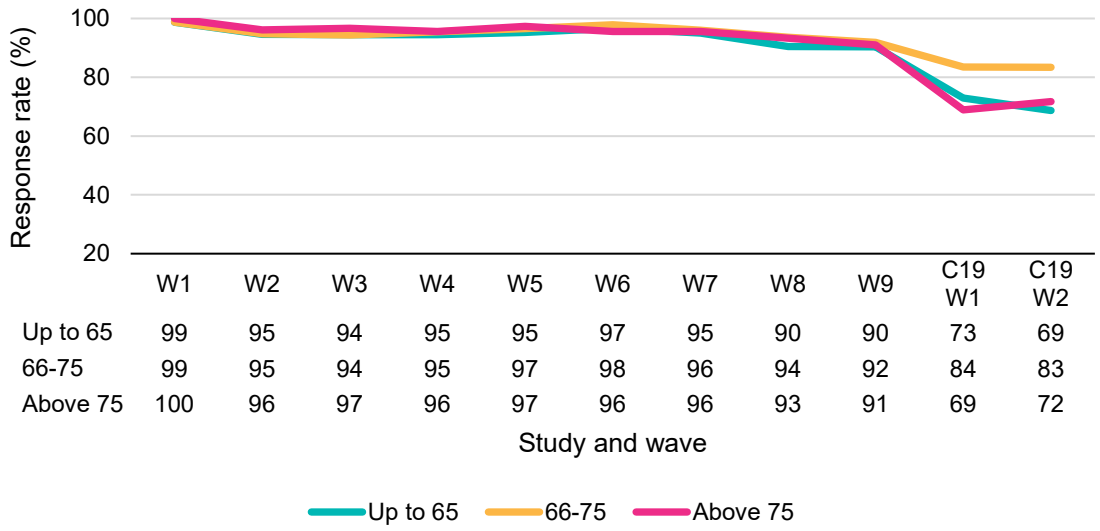
Base: Education below degree 6,036, Education equal or above degree 1,583.

5.2.4 Age groups

While there is a general consensus that younger age groups are less likely to take part in surveys (Watson and Wooden, 2009; Jackle et al. 2015), it is a bit more difficult to understand nonresponse in older age groups, with some studies offering evidence that longitudinal attrition is higher in later life, other studies suggesting it decreases in older age groups, and others founding no evidence in either direction (Watson and Wooden, 2009).

Our findings suggest that the probability of taking part in the ELSA surveys did not significantly vary between CAPI waves across age groups, but sample members aged 76 or above and those younger than 66 years old, were significantly less likely to take part in the ELSA COVID-19 Substudy, compared to those aged between 66 and 75 (Figure 18).

Figure 18: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by age group



	W1	W2	W3	W4	W5	W6	W7	W8	W9	C19 W1	C19 W2
Up to 65	99	95	94	95	95	97	95	90	90	73	69
66-75	99	95	94	95	97	98	96	94	92	84	83
Above 75	100	96	97	96	97	96	96	93	91	69	72

Base: Age up to 65: 6,798 (W1), 6,422 (W2), 6,103 (W3), 5,702 (W4), 5,202 (W5), 4,685 (W6), 3,988 (W7), 3,396 (W8), 2,820 (W9), 2,250 (C19 W1), 1,766 (C19 W2); Age 66-75: 823 (W1), 1,111 (W2), 1,339 (W3), 1,603 (W4), 1,891 (W5), 2,090 (W6), 2,443 (W7), 2,692 (W8), 2,844 (W9), 2,940 (C19 W1), 2,866 (C19 W2); Age above 75: 68 (W1), 156 (W2), 247 (W3), 384 (W4), 596 (W5), 914 (W6), 1,258 (W7), 1,601 (W8), 2,025 (W9), 2,499 (C19 W1), 3,057 (C19 W2).

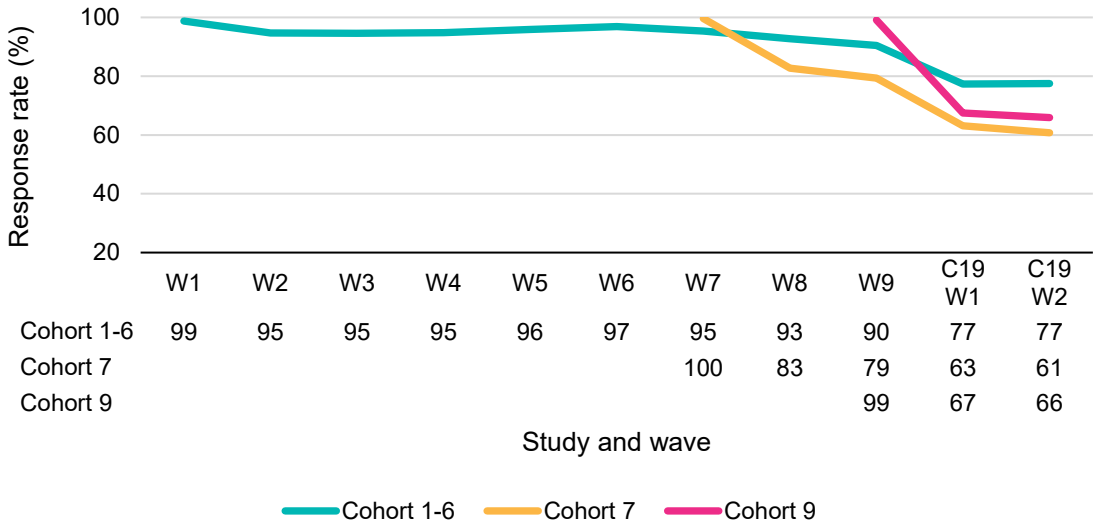
The decline in response rate seen amongst ELSA sample members below the age of 66 might be partly explained by a higher proportion of cases who joined the study either in wave 7 or wave 9 and were included of the most recent refreshment samples. As discussed in the next sub-section, a drop of the response rate in the ELSA COVID-19 Substudy was also found in the most recent study cohorts.

5.2.5 Refreshment samples

We found a steeper decline in the probability of taking part to the Web-CATI study amongst later cohorts (Cohort 7 and 9, who joined the study, respectively, in its seventh and ninth wave; Figure 19). This is aligned to the current literature, suggesting that attrition mostly takes place in the first few waves of the study (Jackle et al. 2015; Lugtig, 2014). Given that this refreshment sample from these two cohorts included the

youngest sample members, the lower response rate seen across those aged up to 65 is likely to be confounded with cohort effects.

Figure 19: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by cohort



	W1	W2	W3	W4	W5	W6	W7	W8	W9	C19 W1	C19 W2
Cohort 1-6	99	95	95	95	96	97	95	93	90	77	77
Cohort 7							100	83	79	63	61
Cohort 9									99	67	66

Base: Cohort 1-6 6,526, Cohort 7 266, Cohort 9 897.

5.2.6 Sex

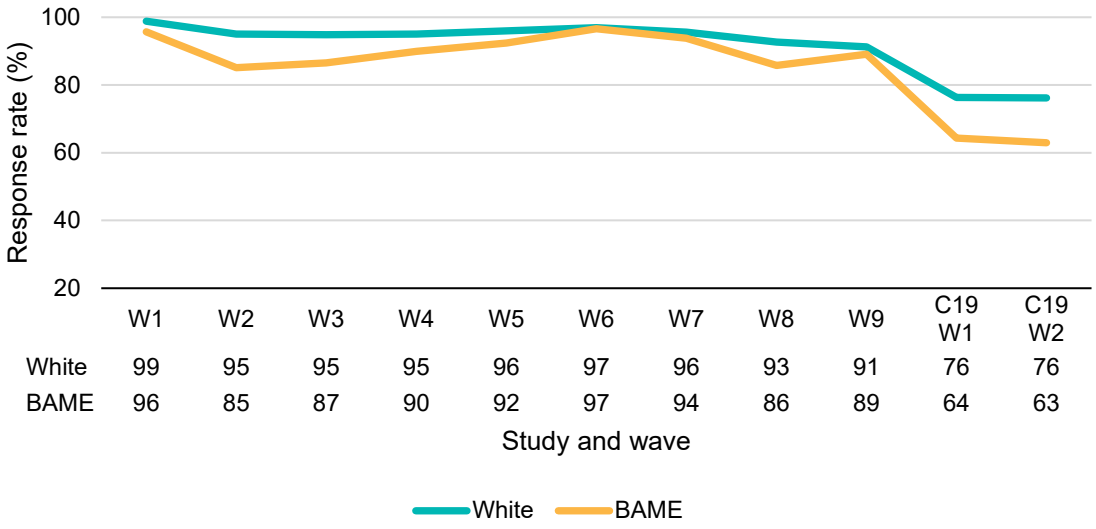
Although studies suggest that women are generally more likely than men to take part in surveys (Watson and Wooden, 2009), we did not find evidence of differences by sex in the participation to the ELSA COVID-19 Substudy.

5.2.7 Ethnicity

Mixed-mode studies are generally associated with lower response rates among ethnic minorities (Feskens et al., 2007; Kappelhof, 2015, Kappelhof, 2017). In the context of a Web-CATI design, the participants' ethnicity can lead to a lower probability of taking part due to due to language difficulties (Young et al., 2006; Becker 2021), especially in the case of first-generation migrants.

ELSA sample members from Black, Asian, Mixed and other ethnic groups were less likely to take part in the ELSA waves, compared to White sample members. However, we did not find evidence of the likelihood varying over time between waves, or the trend become more or less severe in the ELSA COVID-19 Substudy (Figure 20). The peak in response rates seen across sample members from a BAME background in wave 6, 7 and 9 (where it was not different from the response rate of sample members with a White background) reflects the integration of new cohorts in the sample (cohort 6, 7 and 9).

Figure 20: response rate in the ELSA study of core members eligible at the first wave of the ELSA COVID-19 Substudy by ethnic group



Base: White 7,314, BAME 367.

5.3 Nonresponse and the mixed-mode design

Our second research question of this chapter seeks to explore how the implementation of a mixed-mode design has mitigated nonresponse across different subgroups of the ELSA sample compared to a web-only design. While the addition of CATI in a Web-CATI design has increased response rate across all the subgroups, using two modes is expected to have had a greater impact in reducing nonresponse for some sample members, compared to others.

(RQ2) For which subgroups of the ELSA COVID-19 Substudy sample was the Web-CATI design more effective at mitigating Web nonresponse?

For each wave of the ELSA COVID-19 Substudy we computed two different outcome codes. Sample members who took part in the study on Web were considered productive in a Web-only design; sample members who completed the study either on Web or on CATI were considered productive in a mixed-mode design.

The next few sections present the response rates across the two designs and the two waves. Significance testing was carried out in this analysis with multilevel models, assessing if the mixed-mode design option interacted with socio-economic elements and demographic characteristics in predicting a higher likelihood of participating in the study for subgroups of the samples. The models used in this chapter did not include control variables, meaning that the findings presented here should be interpreted considering potential confounded effects of age and cohort.

In our analysis we used the same socio-economic elements and demographics characteristics used in the first part of this chapter to explore nonresponse in the ELSA COVID-19 Substudy compared to the CAPI waves. Our approach aims to determine whether the higher likelihood of nonresponse seen in some groups was mitigated by adding an offline interviewer-based mode (CATI) to the initial online self-administered mode (Web).

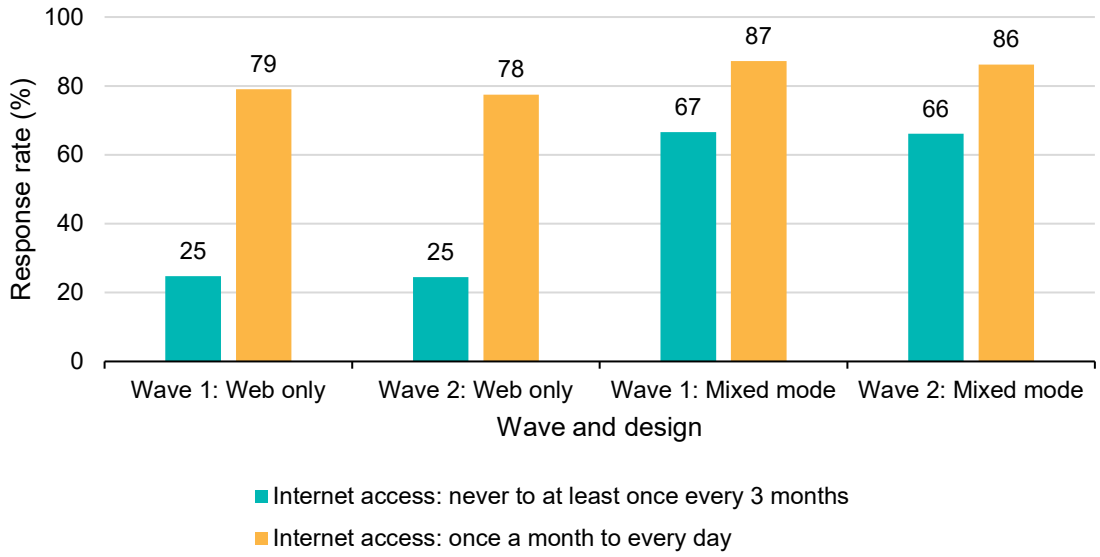
The first limitation of this approach is represented by the fact that sample members became aware during fieldwork of the possibility of a telephone interview, and this might have influenced their participation on Web. Another limitation is represented by the fact that some Web interviews might have been achieved only after the sample members were prompted by the telephone interviewers.

Nevertheless, we accepted this bias in our analysis, on the basis that only a small proportion of respondents were aware of a CATI mode since the beginning of fieldwork (the majority of sample members were informed of possible CATI interviews weeks after Web fieldwork had begun) and that he majority of the Web interviews had been completed before the beginning of CATI fieldwork.

5.3.1 Frequency of internet use

The implementation of a mixed-mode design was successful in achieving a higher participation amongst participants who used the internet less frequently or never (Figure 21). The increase in the probability of taking part when moving from a web-only to a mixed-mode design for this group is higher than the increase observed amongst those who accessed the internet more frequently.

Figure 21: response rate in the ELSA COVID-19 Substudy for different study designs, by frequency of internet access



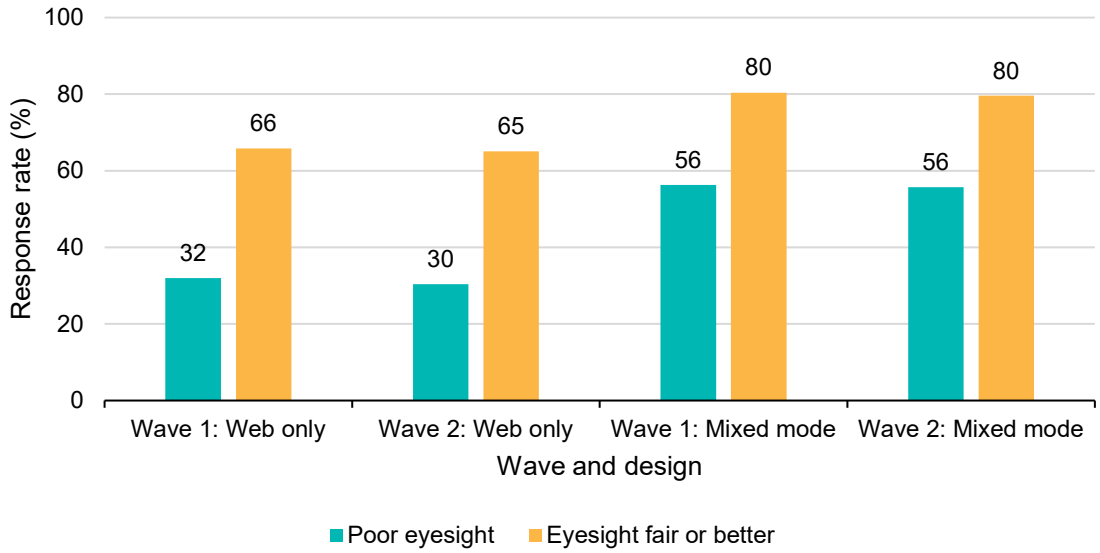
Base: Wave 1 Never to at least once every three months 1,115, Wave 1 Once a month to everyday 5,007, Wave 2 Never to at least once every three months 1,032, Wave 2 Once a month to everyday 4,950.

5.3.2 Self-reported health conditions

The mixed-mode design was associated with higher probability of completion amongst those with poor eyesight, compared to a web-only design (Figure 22).

We did not find statistical evidence to infer that the operationalisation of a mixed-mode design had a greater effect on mitigating nonresponse amongst those who reported poor health (compared to those with better general health) or reported poor hearing (compared to sample members with better hearing).

Figure 22: response rate in the ELSA COVID-19 Substudy for different study designs, by level of eyesight

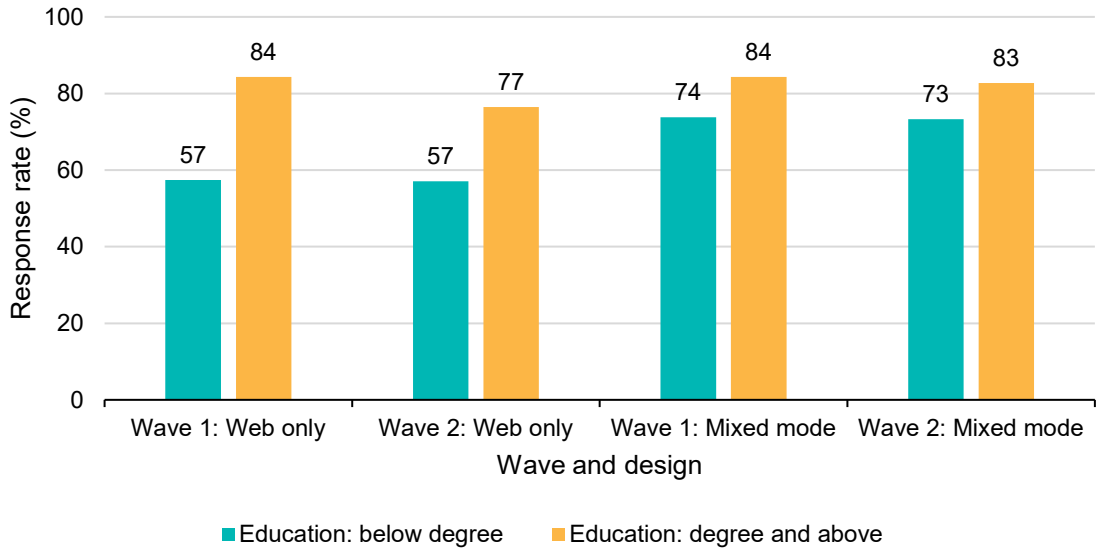


Base: Wave 1 Poor eyesight 206, Wave 1 Eyesight fair or better 6,839, Wave 2 Poor eyesight 194, Wave 2 Eyesight fair or better 6,659.

5.3.3 Highest level of formal education

The mixed-mode design had little impact on mitigating nonresponse compared to a Web-only design amongst those who have a level of education equal or above tertiary education. On the other hand, it increased response rates amongst those with education below degree (Figure 23).

Figure 23: response rate in the ELSA COVID-19 Substudy for different study designs, by level of education

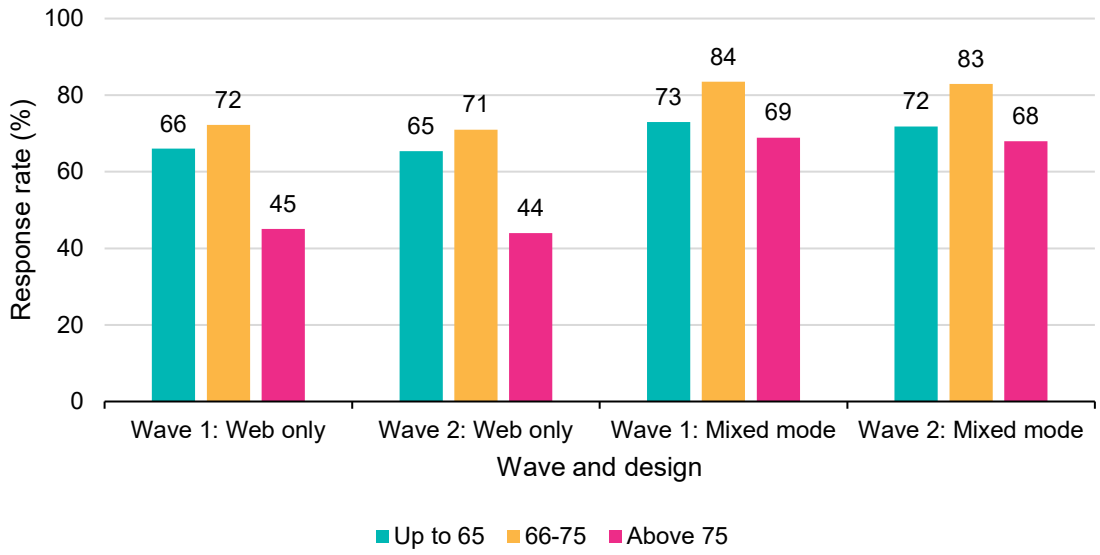


Base: Wave 1 Education below degree 6,036, Wave 1 Education degree and above 1,538, Wave 2 Education below degree 5,833, Wave 2 Education degree and above 1,563.

5.3.4 Age groups

The mixed-mode design mitigated Web nonresponse amongst older sample members. We have found that the mixed-mode design was more likely to have facilitated survey participations for sample members above the age of 75, while its positive effect on mitigating nonresponse was lower for younger sample members (Figure 24).

Figure 24: response rate in the ELSA COVID-19 Substudy for different study designs, by age groups



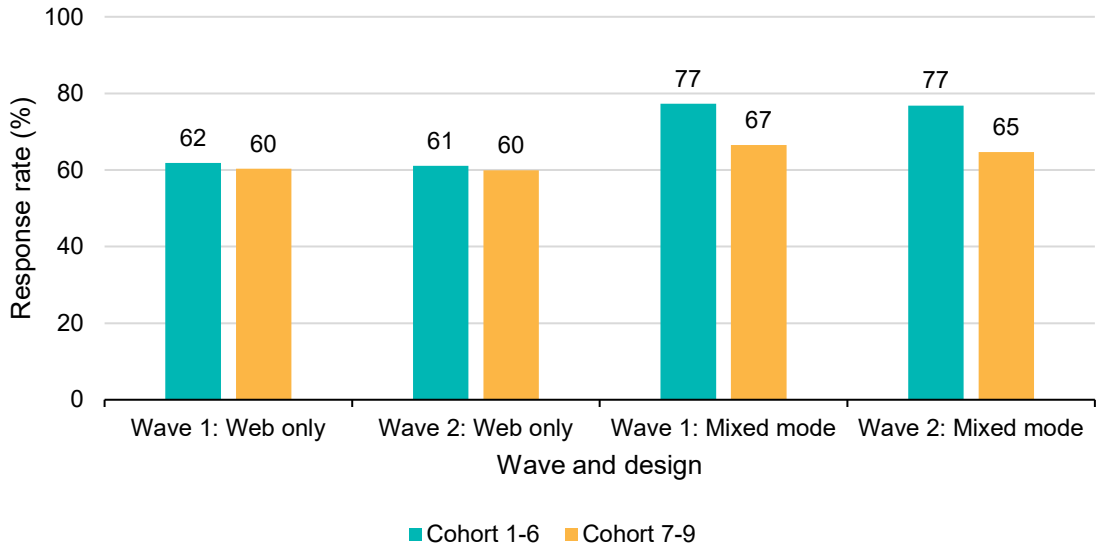
Base: Wave 1 Up to 65 2,250, Wave 1 66-75 2,940, Wave 1 Above 75 2,499, Wave 2 Up to 65 2,238, Wave 2 66-75 2,907, Wave 2 Above 75 2,319.

5.3.5 Refreshment samples

The presence of an additional mode for survey completion was associated with a greater likelihood of completing the study for earlier cohorts, compared to the later ones (Figure 25).

This seems to confirm that the decision of taking part in an additional wave for sample members who recently joined the study comes down to a heuristic cost-opportunity considerations (Goves et al., 2000; Groves and Cooper, 1998), and it is unlikely to be influenced by the interview modes available.

Figure 25: response rate in the ELSA COVID-19 Substudy for different study designs, by sample cohort

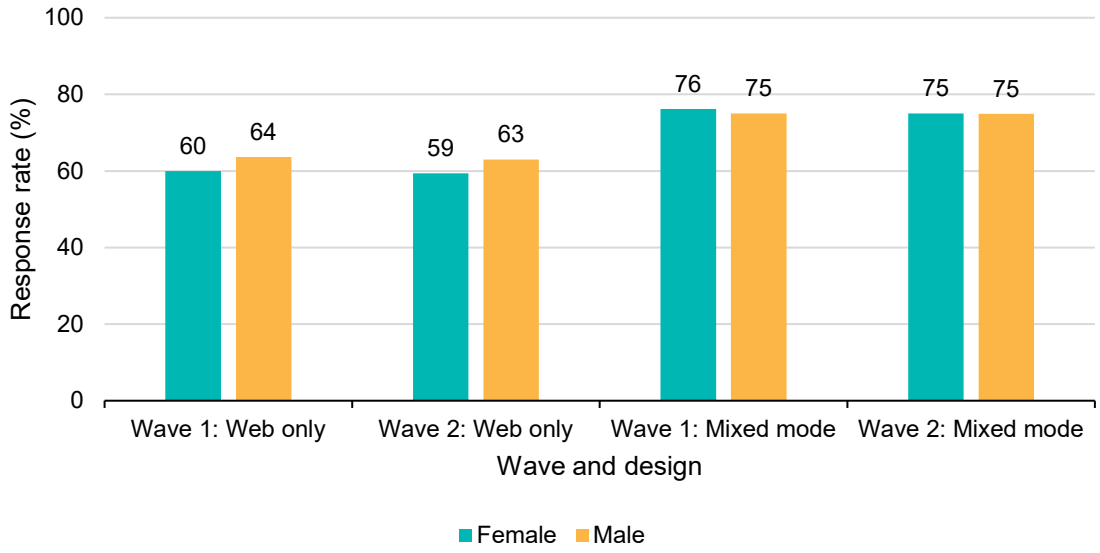


Base: Wave 1 Cohort 1-6 6,526, Wave 1 Cohort 7-9 1,163, Wave 2 Cohort 1-6 6,303, Wave 2 Cohort 7-9 1,161.

5.3.6 Sex

The mixed mode design was found to have had a slightly stronger impact on mitigating response rates for female sample members, compared to the male ones (Figure 26). Response rates across the two waves of the ELSA COVID-19 Substudy were the same by sex under the mixed-mode design, while it would have been higher for male and lower for female with the operationalisation of a Web-only design.

Figure 26: response rate in the ELSA COVID-19 Substudy for different study designs, by sex



Base: Wave 1 Female 4,355, Wave 1 Male 3,334, Wave 2 Female 4,235, Wave 2 Male 3,229.

5.3.7 Ethnicity

We found some statistical evidence to infer that the mixed-mode design had a slightly stronger impact on mitigating nonresponse for White sample members, compared to those with a BAME ethnic background.

However, the impact was so small that the differences in response rates seen between White and BAME ethnic groups were effectively consistent across the two designs.

5.4 Conclusions

There are three broad conclusions that can be drawn on the analysis carried out on this chapter. The first one is that almost all subgroups of the ELSA sample were less likely to take part in the sub-up study, compared to the previous CAPI waves; however, nonresponse was higher in some subgroups than in others. For example, nonresponse was lower in the ELSA COVID-19 Substudy for all the age groups compared to the CAPI design, but the drop in survey participation was steeper amongst sample members aged 75 or above and amongst those below the age of 66.

The second conclusion is that the response rate across subgroups in the substudy was lower compared to the CAPI waves, but not low in absolute terms. With the exceptions of people with poor eyesight, all the subgroups considered in this analysis had a response rate above 60%.

Our third conclusion is that nonresponse, especially in those subgroups who were the least likely to take part in the study, was strongly mitigated by offering an offline interviewer-based mode (CATI) after the initial online self-administered mode (Web). In other terms, the increase in nonresponse observed between the CAPI study and the COVID-19 Substudy would have been greater without offering a CATI mode, especially in some subgroups of the population (such as sample members who do not use the internet frequently, or those who have poor eyesight).

The role of the CATI mode in rebalancing the sample around Web nonresponse bias observed in both waves of the ELSA COVID-19 Substudy is likely to be a consequence of the use of priority groups during for the management of CATI fieldwork. Investing greater time and resources on subgroups of the ELSA sample who were predicted higher nonresponse in Web fieldwork was successful in offering a greater increase in response rate in those groups, rebalancing the samples skewed by Web nonresponse.

Reference list

Addario G., Dangerfield P., Hussey D., Pacchiotti B. and Wood M. (2020), *Adapting fieldwork during the COVID-19 outbreak: A methodological overview of the ELSA COVID-19 Substudy (wave 1)*, ELSA Project, available at: https://www.elsa-project.ac.uk/covid-19-reports.

Becker R. (2021), *The effects of a special sequential mixed-mode design, and reminders, on panellists' participation in a probability-based panel study*, Quality & Quantity, 56: 259–284.

Calderwood, L., Goodman, A., Ploubidis, G., Sakshaug, J. and Silverwood, R. (2019), *An Investigation of Item Non-Response and Measurement Equivalence in a Mixed-Mode Cohort Study: Findings from the Next Steps Age 25 Sweep*, Centre for Longitudinal Studies, Economic and Social Research Council, available at: https://esrc.ukri.org/files/funding/funding-opportunities/an-investigation-of-item-non-response-and-measurement-equivalence-in-a-mixed-mode-cohort-study-findings-from-the-next-steps-age-25-sweep.

Feskens R., Hox J., Lensvelt-Mulders G., Schmeets H. (2007), *Nonresponse Among Ethnic Minorities: A Multivariate Analysis*, Journal of Official Statistics, Vol. 23, No. 3, 2007: 387–408.

Groves, R. M., Singer, E. and Corning, A. (2000), *Leverage-saliency theory of survey participation: description and an illustration*. Public Opinion Quarterly 64 (3): 299–308.

Groves, R. M. and Couper, M. P. (1998), *Nonresponse in Household Interview Surveys*. New York: Wiley.

Jackle A., Lynn P., and Burton J. (2015), *Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-Response*, Survey Research Methods 9(1): 57–70.

Kappelhof, J. (2017), *Survey research and the quality of survey data among ethnic minorities*. In P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. Lyberg, ... B. West (Eds.), Total survey error in practice, (pp. 235–252). Wiley Series in Survey Methodology. Hoboken, NJ: Wiley & Sons.

Kappelhof, J. (2015), *Face-to-Face or Sequential Mixed-Mode Surveys Among Non-Western Minorities in the Netherlands: The Effect of Different Survey Designs on the Possibility of Nonresponse Bias*, Journal of Official Statistics, Vol. 31, No. 1, 2015: 1–30.

Lugtig P. (2014), *Panel Attrition: Separating Stayers, Fast Attriters, Gradual Attriters, and Lurkers*, Sociological Methods and Research 43(4): 699–723.

Pacchiotti B., Hussey D. and Bennett G. (2021), *The dynamics of ageing: The 2018/2019 English Longitudinal Study of Ageing (Wave 9)*, Technical Report, ELSA project, available at: https://www.elsa-project.ac.uk/study-documentation.

Watson N. and Wooden M. (2009), *Identifying Factors Affecting Longitudinal Survey Response*. In P. Lynn (Ed.), Methodology of Longitudinal Surveys, (pp. 157-181). Wiley Series in Survey Methodology. Hoboken, NJ: Wiley & Sons.

Young A., Powers J. and Bell S. (2006), *Attrition in longitudinal studies: who do you lose?*, Australian and New Zealand Journal of Public Health, 2006 vol. 30 no. 4: 353-361.